DISTRIBUTION AGREEMENTS

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations, governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by professor Richard D. Cummings under whose direction it was written, or, in his absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

Signature:

_____ _____

Qian Sun Date

MECHANISMS OF LOCALIZATION OF THE MOLECULAR CHAPERONE

COSMC TO THE ENDOPLASMIC RETICULUM

By

Qian Sun
Master of Science

Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

Richard D. Cummings, Ph.D.
Adviser

Winfield S. Sale, Ph.D.
Committee Member

Victor Faundez, M.D., Ph.D.
Committee Member

Michael Koval, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

Date

MECHANISMS OF LOCALIZATION OF THE MOLECULAR CHAPERONE COSMC TO THE ENDOPLASMIC RETICULUM

By

Qian Sun

M.D., Shandong University School of Medicine, 2003

Adviser: Richard D. Cummings, Ph.D.

An Abstract of

A thesis submitted to the Faculty of the

James T.Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science

Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

2010

ABSTRACT

MECHANISMS OF LOCALIZATION OF THE MOLECULAR CHAPERONE COSMC TO THE ENDOPLASMIC RETICULUM

Qian Sun

Human Cosmc (Core 1 β3-Gal-T-Specific Molecular Chaperone) is encoded by a single exon gene on Xq24 and its cDNA predicts a 318-amino-acid transmembrane glycoprotein with type II topology. Cosmc acts as a specific molecular chaperone for core 1 β3-galactosyltransferase (C1β3Gal-T or T-synthase) and assists in the folding/maturation of this enzyme. Mutation in Cosmc accounts for the defects in T-synthase, which is related to some autoimmune diseases as well as some human cancers. Our previous study showed that Cosmc is primarily localized in the endoplasmic reticulum (ER). Because Cosmc does not have traditional ER retention or retrieval motifs, the ER localization of Cosmc is likely related to other structural features of the protein. Here we explore the potential retention/retrieval mechanism of Cosmc in the ER. In order to address the role of different domains of Cosmc on ER localization, six different chimeric proteins were generated and expressed in COS-7 cells, which allowed us to examine the importance of each domain in the ER localization. By co-localization with intracellular markers, we were able to determine that the transmembrane domain of Cosmc is both necessary and sufficient for its ER localization.

MECHANISMS OF LOCALIZATION OF THE MOLECULAR CHAPERONE COSMC TO THE ENDOPLASMIC RETICULUM

By

Qian Sun

M.D., Shandong University School of Medicine, 2003

Adviser: Richard D. Cummings, Ph.D.

A thesis submitted to the Faculty of the

James T.Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science

Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

2010

ACKNOWLEDGEMENTS

I would like to express my special thanks to my master advisor, Dr. Richard Cummings, Ph.D., for his valuable advice, patience and enthusiasm throughout my graduate school. His wonderful education and scientific guidance made me more precise and mature in science, which will be a treasure for my whole life.

I would also like to thank my committee members, Dr. Win Sale, Dr. Victor Faundez, Dr. Mike Koval and Dr. Anita Corbett for their useful advice and guidance over the past five years.

I would like to dedicate my today's achievements to my dear parents, Jingshun Sun and Yuhua Luo. This dissertation would not have been possible without their unlimited love and support to encourage me throughout these years.

My final acknowledgement goes to the Cummings' lab members. It is such an unforgettable time for me being in this lab for four years. I would like to specially thank Dr. Tongzhong Ju who is an excellent teacher and provided wonderful technical support in my master research.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview of the Secretory Pathway

The application of electron microscopy to cultured cells by Porter led to the unexpected discovery of a major organelle involved in secretion, which was later called the endoplasmic reticulum (ER)(1). Once the proteins synthesized by ribosomes become bound to the rough ER, they enter the ER during their synthesis. Soluble proteins are first localized in the ER lumen, and then sorted into other organelles or secreted from the cell (2). On the other hand, integral membrane proteins are initially inserted into the ER membrane, and subsequently some remain there, but many will be re-localized to the plasma membrane or membranes of the Golgi complex, lysosomes, or endosomes (2,3).

Newly synthesized polypeptides in the membrane and lumen of the ER undergo dozens of possible post-translational modifications, and five of the most common major modifications before they reach their final destinations are: a) formation of disulfide bonds(4); b) proper folding; c) addition and processing of carbohydrates; d) specific proteolytic cleavages; and e) assembly into multimeric proteins (2). Only properly folded and assembled proteins are transported from the ER to the Golgi complex and ultimately to the cell surface or other final destinations(2,5).Unfolded, misfolded, or partly folded proteins are retained in the ER and moved from the ER lumen back into the cytosol where they are degraded by the ubiquitin-mediated proteolytic pathway in proteasomes (2,5).

Most plasma membrane and secretory proteins contain one or more carbohydrate chains. Some glycosylation reactions occur in the lumen of the ER, where they are often involved in quality control of glycoprotein folding and assembly, but most occur in the lumen of the *cis-, medial-, or trans*-Golgi cisternae(2,6). Therefore, the presence of certain types of glycans on proteins and the glycan structures provide useful markers for following their movement from the ER through the Golgi.

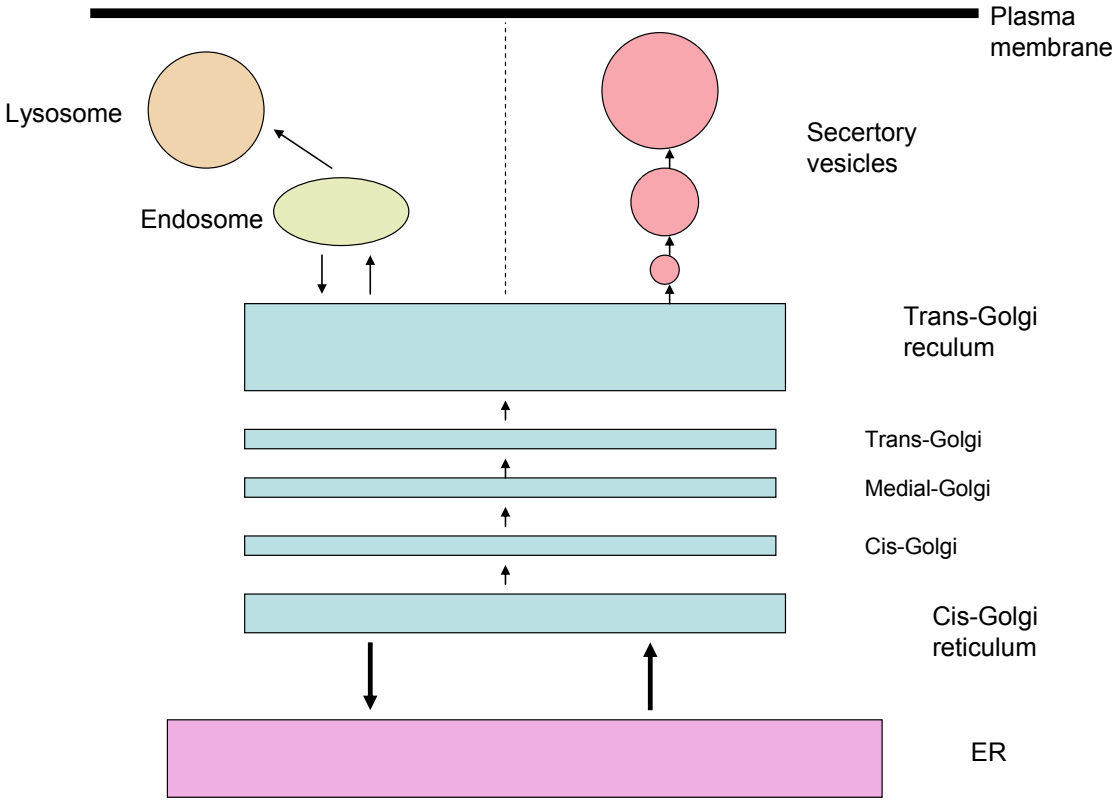


Figure 1. **Schematic diagram of transport steps between membrane-bound organelles in a typical eukaryotic cell.**

1.2 Molecular Chaperones

Although the primary structure of a small protein determines its folding *in vitro* (7), protein folding *in vivo* is usually mediated by accessory proteins known as molecular chaperones (8-10). In molecular biology, chaperones are proteins that assist the non-covalent folding/unfolding and the assembly/disassembly of other macromolecular structures, but do not occur in or with these structures when the latter are performing their normal biological functions. One major function of chaperones is to prevent both newly synthesized polypeptide chains and assembled subunits from aggregating into nonfunctional structures. It is for this reason that many chaperones, but by no means all, are also heat shock proteins, because the tendency to aggregate increases as proteins are denatured by stress or synthesized under stress. Defects in protein folding are associated with several human diseases, such as cystic fibrosis, Alzheimer's disease, spongiform encephalopathies, and familial amyloidotic polyneuropathy (11,12).

In general, there are eight families of molecular chaperones. 1) Small heat shock proteins (hsp25) protect against cellular stress and prevent aggregation in the lens (cataract). 2) Hsp60 system (cpn60, GroEL) ATPase aids in protein folding. 3) The Hsp70 system (DnaK, BiP) ATPase aids in stabilization of extended chains, membrane translocation, and regulation of the heat shock response. 4) Hsp90 ATPase aids in binding and stabilization/ regulation of steroid receptors, protein kinases. 5) Hsp100 (Clp) ATPase is important in thermotolerance, proteolysis, and resolubilization of aggregates. 6) Calnexin and calreticulin are among a group of lectins (glycan-binding proteins) important for glycoprotein maturation in the ER and quality control. 7) Protein

disulfide isomerase (PDI) and peptidyl proline isomerase (PPI) are folding catalysts. 8) Prosequences include alpha-lytic protease and subtilisin (intramolecular chaperones).

The components and mechanisms of action of two major chaperone systems have been well studied (9,13). One system acts on glycoproteins with N-glycans (Asn-linked oligosaccharides) and is dependent on the presence of both monoglucosylated *N*- glycans and unfolded regions on nascent glycoproteins (9,13) and involves the lectin chaperones calnexin and calreticulin (9). The second major ER chaperone system is dependent on the presence of unfolded regions on proteins containing hydrophobic residues, which are recognized by the ER chaperone BiP, an ER Hsp70 family member (9,13). Like all Hsp70 proteins, BiP binds both ADP and ATP, which serve to regulate its binding and release from nascent chains. The hydrolysis of ATP to ADP causes Hsp70 proteins to bind tightly to substrates; the exchange of ATP for ADP induces a conformational change in Hsp70, which in turn causes the release of bound substrates. The ATPase cycle of Hsp70 proteins is both positively and negatively regulated by a number of chaperones and cofactors, including DnaJ, GrpE, Hip, Hop, and Bag-1. Hsp70 proteins are thought to undergo cycles of binding and release from unfolded proteins, with folding occurring during the release cycle(14).

Recent evidence also demonstrates that there are a number of so-called protein-specific chaperones, whose major function is to assist in the folding and maturation of a single specific protein (8). Some of these are required for protein folding in the ER and some also serve as so-called escorts, outfitters, and guide proteins for protein exit to the Golgi (15). Examples include HSP47(16-18), SCAP(19), RAP, and MESD. RAP

(receptor-associated protein) is a soluble ER molecular chaperone with a HDEL-type retention motif that binds to LDL-related proteins (LRPs) in the ER and prevents their binding to endogenous ligands and subsequent degradation and escorts LRP to the Golgi (20,21). Mesodermal development protein (MESD) is an ER chaperone that promotes disaggregation and proper membrane localization of the WNT coreceptors LRP5 and LRP6. In the absence of sufficient MESD, LRP5 and LRP6 form intermolecular aggregates due to improper folding and disulfide bond formation. Like RAP, MESD may have a related function to biosynthesis of LRP family members (22).

1.3 Background on protein glycosylation

Glycosylation is the process of addition of sugar to proteins and lipids, which is an enzyme-directed and site-specific process. There are two common types of glycosylation: N-linked glycosylation to the amide nitrogen of asparagine side chains (N-glycans) and O-linked glycosylation to the hydroxyl oxygen of serine and threonine side chains (O-glycans). However, over 25 different types of linkages and structures are known between glycans and amino acids in proteins. N-glycosylation of nascent polypeptides occurs co-translationally, while O-glycosylation (mucin-type) occurs post-translationally. After immature glycoproteins are generated in the RER lumen, they are subsequently transported into Golgi apparatus and modified by removal of sugar from N-glycans and addition of sugars to both N- and O-glycans.

All N-glycans are synthesized in the rough ER with addition of a large preformed oligosaccharide precursor (2,23). Formation of this precursor oligosaccharide begins with assembly of a ubiquitous 14-residue high-mannose precursor on dolichol, which is a lipid

in the membrane of the ER. Then this precursor is transferred to the specific asparagine residues of nascent polypeptide in the ER (2,23). After adding this preformed oligosaccharide in the ER, three glucose residues and one mannose residue will be sequentially removed by specific glycosidases on the fully-folded proteins. After this processing, newly synthesized glycoproteins enter the *cis-, medial-,* and *trans*-Golgi cisternae, where other mannose residues are removed and additional sugars added, one at a time from nucleotide sugar donors by various enzymes, to yield a finished N-glycan. N-glycosylation is an important process in the cells, and is especially important in glycoprotein biosynthesis and correct protein folding (2,23). The majority of non-lectin like chaperones modulates the folding of the newly-synthesized glycoproteins through their non-specific interaction with hydrophobic domains exposed in immature glycoproteins polypeptide motifs. However, there is a host of lectin chaperones that recognize glycans and/or protein features on newly synthesized glycoproteins. Two of the most well studied are, calnexin and calreticulin, which interact with the monoglucosylated N-glycans within the nascent glycoproteins during the early steps of the N-glycan processing in the ER (24). Glucose is removed by glucosidase II and then glucose may be readded by protein:UDP-Glc glucosyltransferase that recognizes unfolded glycoproteins in a cycle until the glycoproteins are folded. Additional mannose-binding proteins in the ER, such as OS-9 (25) and EDEM (26), also serve as chaperones in the quality control system. Incorrect glycoprotein folding can lead to export of the glycoprotein from the ER to the cytoplasm, where the N-glycans are removed and the protein is ubiquitinated and degraded in the proteasome(27).

O-glycan biosynthesis occurs in the Golgi apparatus and is relatively simpler than N-glycosylation in that a lipid-linked oligosaccharide precursor for transfer to protein is not required(28). Nucleotide diphosphate or monophosphate sugars are the immediate precursors used in the formation of O-glycans(29). These monosaccharides are added one at a time, and each transfer is catalyzed by a different glycosyltransferase, which is specific for the acceptor glycan or aglycone (protein or lipid) and donor nucleotide(29). The initiating event in mucin-type O-glycan biosynthesis is the addition of the monosaccharide GalNAc (from UDP-GalNAc) to serine and threonine residues catalyzed by a family of polypeptide α-GalNAc transferases (ppGalNAcTs)(29). Afterward, this GalNAcα1-Ser/Thr (the Tn antigen) is modified by addition of other sugars, most notably galactose to form more complex-type O-glycans.

While glycoproteins with a few mucin-type O-glycans are common in glycoproteins on all animal cell surfaces, mucins are a special class of heavily O-glycosylated glycoproteins on the cell surface as transmembrane glycoproteins or in mucous secretions(30). The clustering of O-glycans on mucins is due in part to the presence of repeated peptide stretches called “variable number of tandem repeat” (VNTR) regions that are rich in serine and threonine and proline residues(30). In the secreted mucins of the respiratory, gastrointestinal, and genitourinary tracts, as well as eyes, the O-glycans are essential for their ability to hydrate and protect the underlying epithelium(30). Because the O-glycans are hydrophilic and usually negatively charged, they promote binding of water and salts and are major contributors to the viscosity and adhesiveness of mucus, which forms a physical barrier between lumen and epithelium. However, in diseases such as cystic fibrosis, the abnormally high viscosity of mucus

causes airway obstruction(31). In addition, the O-glycans on mucins are important ligands for a wide variety of glycan-binding proteins in all connective tissue and epithelial surfaces, as well as being receptor for pathogens (viruses and bacteria)(32).

Overexpression of mucins in carcinomas has been described for many years. So far, there are two mechanisms to address how abnormal mucins can lead to cancer. One abnormal feature of carcinoma mucins is loss of correct topology(33,34). In the normal polarized epithelial cells, mucins are expressed exclusively on the apical side of cells, toward the lumen of a hollow organ. However, in the carcinoma cells, membrane-bound mucins are expressed on all aspects of the cells, and soluble mucins can even enter the extracellular space and body fluids such as the blood plasma(33,34). These two abnormal forms of mucins can play a critical pathophysiological role in malignancy(34). Another abnormal feature is incomplete glycosylation, which results in the expression of the Tn antigen, sialylated Tn antigen, or T antigen. It is noted that excessive Tn and sialyl Tn antigen are present on tumor cells(33).

1.4 Background on Glycosyltransferases

Glycosyltransferases are a group of very prevalent enzymes, mostly found in the ER and Golgi apparatus, which represent 1-2% of the genome. Most Golgi glycosyltransferases are type II transmembrane proteins with a short N-terminal cytoplasmic domain, a single transmembrane domain (TMD), and a large luminal C-terminal catalytic domain (35-37). Like other type II membrane proteins, they lack a cleavable signal sequence. Although many glycosyltransferases are concentrated in subcompartments of the Golgi, there is often overlap in the distribution of enzymes between subcompartments. The basis for

Golgi targeting of glycosyltransferases is still poorly understood, and only a few potential mechanisms responsible for their localization in the Golgi have been proposed. For instance, the TMD may participate in Golgi retention either by facilitating protein-protein interaction or by physical mechanisms (38-40). The finding that the length of the TMD appears to be more important than its amino acid composition in the Golgi retention of fusion proteins has favored the second hypothesis that TMD length is a general feature for protein retention in the Golgi from mammalian cells (36,39,41,42), such as the trans Golgi resident enzyme β-1,4-galactosyltransferase. However the luminal domain and/or the cytosolic tail also appear to be important for Golgi residency of some glycosyltransferases (36,43). Recently, a novel ER export motif, [RK](X)[RK], that is cytoplasmically located proximal to the TMD Golgi resident glycosyltransferases, appears to be required for the exit of some enzymes from the ER (44), although this does not explain their residence in the Golgi. ER export may result from binding of Sar1-GTP to ER membranes, probably through a hydrophobic motif exposed at its N-terminus, that induces the formation of ER exit sites and budding of COPII vesicles (45). However, many glycosyltransferases lack this motif. Clearly, much remains to be learned about the mechanisms of Golgi localization of glycosyltransferases. Another poorly understood aspect of glycosyltransferase biosynthesis is the possible regulation due to folding/maturation within the endoplasmic reticulum, since most glycosyltransferases are disulfide-bonded dimers.

1.5 Background of T-synthase

Core 1 β3galactosyltransferase (core 1 β3GalT, T-synthase) specifically transfers Gal to GalNAcα1-Ser/Thr (Tn antigen) using the donor UDP-Gal to synthesize the core 1 O-glycan Galβ1-3GalNAcα1-Ser/Thr (T antigen). The T antigen is a key precursor for most common mucin-type O-glycans on both of transmembrane and secreted glycoproteins in vertebrates and invertebrates. Therefore, the overall pathway of mucin glycosylation is regulated by this critical branch point T-synthase. T-synthase is a disulfide bonded, homodimeric type II transmembrane protein, encoded by the human *T-synthase* gene on chromosome 7 (7p14-p13) (29,46). The enzyme is developmentally important and disruption of the *T-synthase* in mice is embryonic lethal due to the impaired angiogenesis (47,48). Interestingly, disruption of T-synthase in hematopoietic and endothelial cells in mice results in misconnection of blood and lymphatic vessels, indicating the important role of O-glycans in lymphoangiogenesis(47,48). Loss activity of T-synthase will lead to a accumulation of Tn and STn(48). A number of autoimmune diseases, including IgA nephropathy, Tn syndrome, and Henoch-Schönlein purpura are related to abnormal expressions of Tn and STn. Tn and STn are also recognized as tumor antigens as seen in many human cancers, such as carcinomas of breast, bladder, cervix, pancreas, ovary and colon and associated with tumor's poor prognosis (35,49-53). Understanding of molecular mechanism for how T-synthase activity is regulated will certainly aid us in developing new therapeutic methods against these diseases.

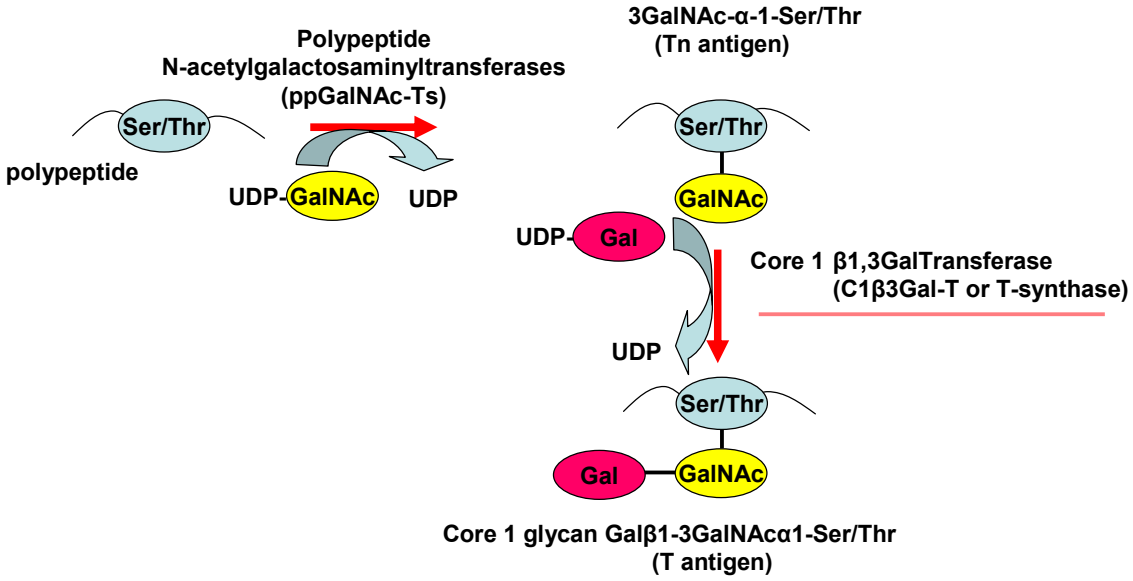


Figure 2. **Schemic diagram of catalytic function of T-synthase in regulating biosynthesis of T antigen.**

1.6 Background on Cosmc, a specific molecular chaperone for T-synthase

Cells of hematopoietic origin in patients with Tn syndrome and many types of tumor cells lack T-synthase activity, although they have the T-synthase transcript. In such cases cells express the Tn antigen, indicating a lack of branch point enzyme activity of the T-synthase or other enzymes. Jurkat cells, which is a human T cell lymphoblastic-like cell line, was established from the peripheral blood of a 14-year-old boy with acute lymphoblastic leukemia (ALL) at first relapse in 1976(54). Jurkat cells lack T-synthase activity and synthesize Tn and sialyl Tn antigens on all mucin-type O-glycans (55). Since the cells lack T-synthase activity, it was assumed that the cells had either a mutation in the *T-synthase* or had transcriptional defects in T-synthase expression. However, after

further studies, no defects in the coding sequence or expression of mRNA for T-synthase were found in Jurkat cells. In order to explore the mechanism for regulation of T-synthase activity, our laboratory partially purified T-synthase from rat liver and two proteins were identified by sequencing the N-terminus of this purified material (55). One sequence belonged to murine T-synthase, but the other was derived from an unknown protein. Later this protein was designated as Cosmc (Core 1 β3-Gal-T-Specific Molecular Chaperone) (55).

Cosms is a unique and apparently client-specific chaperone for the T-synthase, and is required for formation of active T-synthase. Several important findings led to the conclusions of our laboratory that Cosmc functions as a molecular chaperone for T-synthase. (**#1**) Many types of human cancer cells, including human T-lymphoid Jurkat, colon carcinoma LSC, LS174T, and human melanoma LOX cells, which lack T-synthase activity and express Tn/STn antigens, either contain mutated *Cosmc* or have lost the *Cosmc* transcript (55,56); (**#2)** expression of wild type *Cosmc* in Jurkat, LSC, and LOX cells complements T-synthase activity and abolishes expression of Tn/STn antigen (55,56); (**#3**) Tn Syndrome, which is characterized by expression of Tn/STn antigens on blood cells of all lineages, is caused by a somatic mutation in *Cosmc* in multipotential hematopoietic stem cells (55,56); (**#4**) expression of T-synthase in insect cells, which lack a *Cosmc* ortholog, results in an inactive T-synthase protein; co-expression of Cosmc and T-synthase results in an active T-synthase (55,56); (**#5**) Cosmc is an ER-localized molecular chaperone preventing aggregation/proteasomal degradation of T-synthase (55,56); (**#6**) Cosmc, but not the T-synthase, has an ATP-dependent binding property(55,56); (**#7**) disruption of *Cosmc* in mice results in an embryonic lethality and

Tn expression *in vivo* (55,56). A model of Cosmc function, consistent with the above evidence, is shown in **Figure 3**.

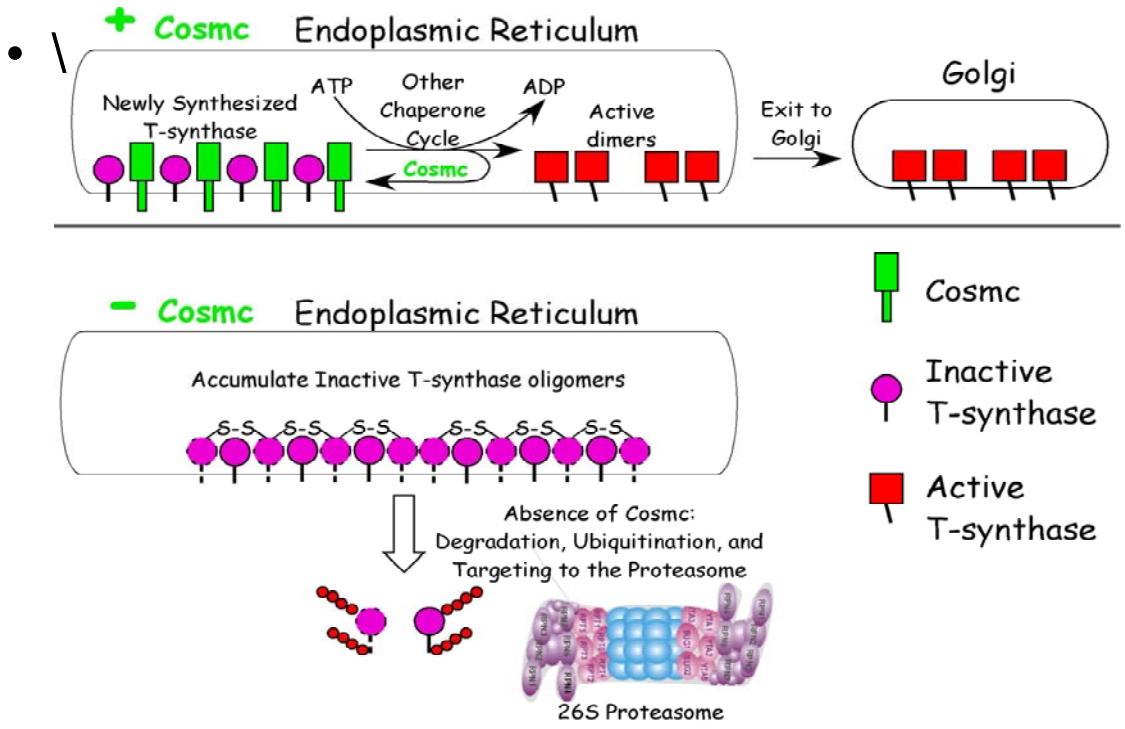


Figure 3. **Working model for Cosmc function as an ER-localized molecular chaperone preventing aggregation/proteasomal degradation of T-synthase. When Cosmc is present in ER, it will direct the newly-synthesized T-synthase to form active dimmers, which will be transported into Golgi apparatus. On the other hand, when Cosmc is absent in ER, newly-synthesized T-synthase will accumulate into inactive oligomers, which will be degraded by proteasome.**

1.7 Background on ER retention/retrieval of proteins

The major question that I addressed in my thesis project was the mechanism of Cosmc localization to the ER. Therefore, it is important to discuss what is known about general ER retention/retrieval mechanisms. The ER is an organelle containing a large number of newly-synthesized proteins, as well as proteins that are maintained in the ER responsible

for posttranslational modifications. These maintained proteins are called ER resident proteins, and their localization to the ER depends on their primary amino acid sequences. Two types of signals have been reported to be involved in ER localization: 1) a static retention signal can keep proteins at a particular location within ER (35,51-53); 2) a dynamic retention signal, also known as retrieval signal, can return proteins to the ER which have escaped to the post-ER compartments. These two types of signals are not exclusive and can function either in parallel or in combination(57-60).

So far, there are several ER retential/retrieval mechanisms, but two of the well-characterized ER-retrieval mechanisms are the KDEL and di-lysine motifs. The KDEL mechanism involves the KDEL tetrapeptide (HDEL in yeast) at the extreme carboxyl terminus of soluble ER resident proteins (57,59,60). PDI, luminal Hsc70, and many other ER-resident proteins, including the lectin chaperone calreticulin, have this KDEL motif. Many experiments have shown that this KDEL tetrapeptide is both necessary and sufficient for retention in the ER. For example, when a protein that normally contains the KDEL at its C-terminus is mutated to lack these four residues, it is secreted out of the ER (2,61). However, if a protein that does not contain these four amino acids is engineered to contain KDEL, it is retained in the ER (2,61). Proteins with the KDEL recognition sequence, which have escaped to the cis-Golgi complex interact with the KDEL receptor (Erd2p in yeast), to return to the ER (2,62). This mechanism has been supported by several experiments. For instance, most KDEL receptors are localized to the membrane of the cis-Golgi network as well as membranes of small transport vesicles shuttling between the ER and the cis-Golgi (2,63). Interactions of proteins with the KDEL receptor returns the protein to the ER via COPI vesicles (2,62). Moreover, ER resident proteins

with the KDEL recognition sequence contain oligosaccharides, which are modified by enzymes found only in the cis-Golgi. Therefore, these proteins must have been transported to the cis-Golgi at some time.

The second mechanism consists of the double-lysine motif [K(X)KXX, where X is any amino acid] in the cytoplasmic domain of transmembrane type I ER resident proteins as well as the double-arginine motif [RR] within the first five amino-terminal residues of transmembrane type II ER resident proteins (57,59,60,64). Similar to the studies about the KDEL noted above, both KK- and RR-containing proteins obtain Golgi modifications and co-localize in the post-ER vesicular structures, which mean that these ER resident proteins have escaped from the ER and later are returned to the ER. Two possible candidates for the retrieval pathway are microtubules (65,66) and the coatomer complex (67,68). Interestingly, it has been reported that in cells treated with nocodazole, a drug that disrupts microtubules, the treated cells have increased Golgi levels of KK-containing reporter proteins, which indicated that an intact microtubule normally serves as a retention matrix for ER membrane proteins (65,66,69). Masking of the KK-motif prevents coatomer binding, which limits the exposure of KK-proteins to the Golgi-localized transferase (69). On the other hand, studies also show that the KK motif-containing proteins are returned from the Golgi to the ER in a coatomer-dependent manner (69).

Another mechanism of ER retrieval is seen in the yeast protein Rer1p (retrieval to ER 1 protein). RER1p localizes several transmembrane proteins such as Sec12p (which is a type II transmembrane protein) and Sec71p to the ER by retrieving them from the

Golgi apparatus. Rer1p is mainly localized to the early region of the Golgi apparatus at a steady state, but it is actively cycling between the Golgi apparatus and the *ER in vivo* in a COP II and Sec18-dependent manner (70). While the precise mechanisms are unclear, Rer1p recognizes a retrieval signal contained in the transmembrane domain of Sec12p (70,71). Interestingly, the transmembrane domain of Sec12p is sufficient to confer Rer1p-dependent ER retention to other membrane proteins(72). Also, there are unknown and novel ER retention or retrieval motifs in the cytoplasmic N-terminal domains of ABCG5 and ABCG8, which are transport proteins for sterols and xenobiotics (73).

The cysteine residues in ER resident proteins may also be important for ER retention. For example, Ero1α and Ero1β are ER resident proteins that transfer oxidative equivalents to protein disulfide isomerase (PDI), the enzyme that oxidizes cargo proteins destined for export from the ER. Neither Ero1α nor Ero1β contains known ER localization motifs, but it was shown that covalent interactions of these proteins with PDI or ERp44, in a cystine-dependent manner, is essential for ER retential of Ero1α and Ero1β (74) .

However, while much is being learned about the mechanisms of ER retention, the total number of retention pathways and the specific mechanism these retrieval signals are not well understood, and novel mechanisms for retention/retrieval probably await discovery. For example, some retrieval signals can also be found in ER non-resident proteins, such as type I transmembrane protein ERGIC53, and type II transmembrane protein p63 (57,75,76). Furthermore, removal of the K(X)KXX motif from an ER resident protein UDP glucuronosyltransferase does not cause it to lose its ER retention,

but the protein lacks its Golgi modifications (57,65). According to general concepts in the field, retention signals probably exist in the ER resident proteins and function in a retrieval signal-independent way. In mammalian cells, there have been many studies to support a role for the transmembrane domain and part of the flanking regions in the retention of Golgi enzymes. The mechanism, as discussed above, is proposed to be a result of the difference in thickness of the lipid bilayer between the plasma membrane and the Golgi membrane (77). Another model postulates that the transmembrane domain interacts with other Golgi proteins to form large oligomers, which prevents them from entering transport vesicles (57).

Unfortunately, the role of the TMD in the retention of ER resident proteins in mammalian cells is less clear. The role of the TMD in the localization of a few yeast ER proteins has been studied, such as a yeast ER t-SNARE Ufe1p and another yeast ER protein Sec12p (72). However, the retential signal is not only determined by the length of the TMD but also is dependent on the amino acid sequence (69). Thus, many questions remain about the mechanisms of ER retention for ER resident proteins. Our studies on the ER retention of the molecular chaperone Cosmc provides new insights on the pathways of ER retention and suggest further studies are needed on precise roles of the amino acid sequences within the TMD of ER resident transmembrane proteins.

Chapter 2

ER-localization Determinants of Cosmc

2.1 Introduction

The molecular basis of specific integral membrane protein retention in different organelles along the secretory pathway of eukaryotic cells is not well understood. The basis for sorting and transport of integral membrane proteins to their different intracellular destinations is of central importance for maintenance of cell integrity. In the past, many studies on the targeting and retention of resident Golgi proteins have been performed, but little is known about ER resident proteins.

After translocation into the ER, it is thought that proteins move by default towards the plasma membrane unless certain specific structure motifs are present, which include a retention signal or a retrieval signal, determining the retention in specific organelles(78). The availability of cDNAs encoding different ER proteins allowed the systematic study of the amino acid composition and structure of the different domains and their role in targeting and ER retention(61,64,79). The generation of chimeric proteins by manipulation of these cDNAs showed that the presence of the tetrapeptides KDEL or HDEL, or double lysine KK of eukaryotic proteins is sufficient for retention and/or retrieval to the ER(61,64,79).

Recently, the use of protein chimeras has yielded further indications that the specific domains of different type II transmembrane Golgi enzymes are involved in Golgi retention(80). The transmembrane domain (TMD) of β1,4 galactosyltransferase, a trans-

Golgi resident enzyme, is sufficient to retain a chimera of an endocytic compartment resident protein in the trans-cisternae of the Golgi stack (81). The TMD of another trans-Golgi resident enzyme, α2,6 sialyltransferase, also specifies Golgi retention of a fused marker protein, the normally secreted chicken lysozyme (80). The TMD of the medial-Golgi resident enzyme, N-acetylglycosaminytransferase I was also shown to mediate the transport of a reporter gene product to the Golgi complex(82).

However, the role of the TMD and other domains of resident ER transmembrane proteins has not been well studied. So far, the general idea about the role of the TMD in the ER localization is focused on its length or hydrophobicity of the TMD(83). Short TMD with hydrophobic residues may promote the ER targeting of those resident proteins(83). In addition, the polar residues within the TMD might be the retrieval signal for some ER resident proteins which can be returned from the Golgi to the ER in a receptor-dependent way(70).

A novel ER-localized molecular chaperone termed Cosmc was recently discovered and shown to assist in the folding and prevention of oligomerization of the key enzyme involved in mucin-type O-glycosylation, the core 1 β1,3 galactosyltransferase (T-synthase). The T-synthase is a Golgi enzyme that adds galactose residues donated from UDP-Gal to glycoproteins entering the Golgi that have the mucin-type sequence GalNAcα1-Ser/Thr to generate the core 1 disaccharide Galβ1-3GalNAcα1-Ser/Thr. Cosmc is a type II transmembrane protein (~35 kDa) with a short cytoplasmic N-terminal domain, and a large ER luminal domain that is not glycosylated and is the functional chaperone domain.

ClustalW Formatted Alignments

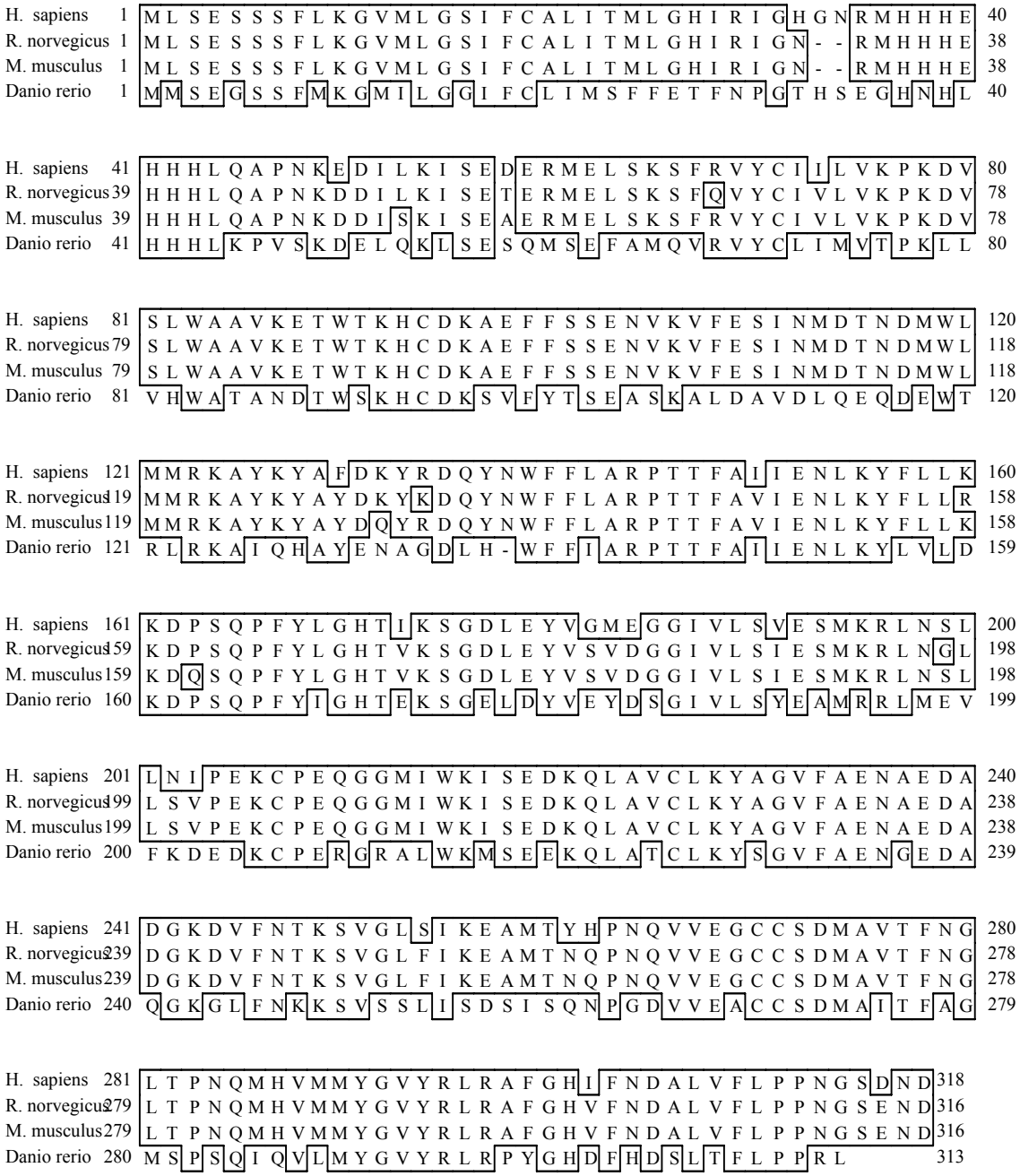


Figure 4. Comparison of amino acid sequences of wild type Cosmc from different species.

There are orthologs for Cosmc in vertebrates from human to zebra fish, but not in invertebrates (56). Human and primate Cosmc have only one single amino acid difference at V191I, and both proteins have 318 amino acids (56). In contrast to human and primate Cosmc, rodent Cosmc contains 316 amino acids and has a two-amino-acid gap between positions 33 and 34, which are located at the beginning of the functionally luminal domain of Cosmc(56). However, the rodent Cosmc has >95% identity to human Cosmc. Frog Cosmc is a 317-amino-acid protein, which has one amino acid gap in the sequence at position 33 compared with human Cosmc. Zebrafish Cosmc lacks 3 amino acids at the C-terminal domain compared to human Cosmc, ending at position 315 (56). Mouse Cosmc can act as a molecular chaperone for both human and mouse T-synthase, but the functions of other Cosmc orthologies have not been tested (56).

The mechanism of retention of Cosmc in the ER is not known. The primary amino acid sequence of Cosmc lacks traditional retention and retrieval signals such as KDEL motif, K(X)KXX motif, or RR motif. ER-localization of Cosmc is essential for its chaperone function for T-synthase since T-synthase is a secretory pathway protein. Soluble secreted form of Cosmc has a little function, while a recombinant soluble form Cosmc engineered to contain a KDEL-tag functions as efficiently as wild type Cosmc (56). Furthermore, a start codon mutation of *Cosmc* identified in a Tn syndrome patient results in unfunctional Cosmc, although it could be translated into a polypeptide lacking the cytoplasmic domain and part of the TMD by using another ATG codon in frame at position 13 (amino acid sequence number) as a start, which is presumably mislocalized in cells (56). Thus, there is a need to define the mechanism of ER localization for Cosmc, which may help to shed light on general ER retention mechanisms. Furthermore,

defining of the determinants for Cosmc ER-localization will aid in fully understanding the functions of Cosmc and its role in human disease processes.

The TMD of Cosmc has 18 amino acids -G-V-M-L-G-S-I-F-C-A-L-I-T-M-L-G-H-I- and it is conserved in many species. Our previous studies have shown that human Cosmc mainly localizes in ER, whereas recombinant, soluble recombinant Cosmc was secreted into the culture media. The secretion of this soluble form of Cosmc and the ER retention of the engineered KDEL-tagged soluble Cosmc (56) indicates that the TMD of Cosmc may be important for its ER-localization. Many previous studies have shown that the TMD could be a determinant for ER retention or retrieval signals(70,83,84). To aid in identifying the determinants that localize Cosmc to the ER, we generated molecular a variety of molecular constructs to test the role of the TMD and other domains on its ER retention. Our approach involved the generation of chimeras by using another type II transmembrane protein, the transferrin receptor (TfR), which has 760 amino acids, but is localized to the plasma membrane and the endosomal vesicles in all animal cells(85). The TfR is ideal for these studies, since it is a well-studied protein and is also a type II transmembrane protein similar to Cosmc. Six different chimeric constructs were generated by recombinant DNA techniques.

Each construct was transiently expressed in COS-7 cells and the intracellular destination of the recombinant protein was followed by indirect immunofluorescence with a monoclonal antibody against a HPC4 tag, which has been added to the C-terminus of each construct, and by subcellular fractionation/Western Blotting method. The HPC4 epitope is a 12 amino acid determinant in the protein C activation peptide region and is

recognized by the Ca^{2+}-dependent HPC4 monoclonal antibody (86,87). It should be noted that many studies on protein localization in the ER and Golgi rely on changes in protein glycosylation dependent on the localization or retrieval. However, Cosmc is not a glycoprotein and thus it is not possible to rely on such indirect approaches for gauging localization. Thus, our approach utilized two key observations. We examined subcellular localization of Cosmc constructs and other markers by confocal immunomicroscopy using intact versus permeabilized cells as well as subcellular fractionation using sucrose gradients of membranes and Western blotting of the recombinant proteins versus other markers.

2.2 Materials and methods

2.2.1 Reagents

African Green Monkey SV40-transfected kidney fibroblast cell line COS7 was obtained from American Type Culture Collection. Restriction enzymes were obtained from New England Biolabs, Inc. FuGENE6 and Taq DNA polymerase was obtained from Roche. TNM-FH and EX-Cell 405 media, rabbit anti–human calnexin mAb (IgG1), and mouse anti-human KDEL (GRP78) mAb (IgG1) were purchased from BD Biosciences. Rabbit anti-human giantin mAb (IgG1) was purchased from Abcam. Rabbit anti–human calnexin antiserum and mouse anti-KDEL (GRP78 and GRP94) mAb (10C3) were purchased from Assay Designs. Alexa Fluor–labeled secondary antibodies were purchased from Invitrogen. Proteasome inhibitors MG-132 and rabbit anti–human proteasome 20S α-type1 subunit (IgG) were purchased from EMD. Vector pcDNA3.1(+), PCR TOPO4 cloning kit and SDS-PAGE gels were obtained from Invitrogen. QIAquick

Gel Extraction kits were obtained from QIAGEN. Chemiluminescent Substrate and BCA protein assay kit were purchased from Thermo Fisher Scientific.

2.2.2 Preparation of expression constructs

A construct encoding C-terminal HPC4-tagged Cosmc (Cosmc-HPC4) was made by introducing the HPC4-epitope (KGDILRPDVQDE) into wild type Cosmc at its C terminus by PCR. The product was cloned into PCR3.1. The insert was cut with BamHI (partially)–XbaI and cloned into pcDNA3.1(+). A construct encoding C-terminal HPC4-tagged TfR (TfR-HPC4) was made using a similar strategy to Cosmc-HPC4. The HPC4 epitope tag was introduced into the C terminus of TfR by PCR. The product was cloned into PCR3.1. The insert was cut with BamHI–XbaI and cloned into pcDNA3.1(+).The PCR primers are listed in Table 1.

Table 1. PCR primers used for making chimeric constructs.

Constructs		Forward Primer		Reverse primer
Cosmc-HPC4	I	5’-CAATGAAAAGACTTAACAG CCT-3’	II	5’-GCTCTAGACTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCGTCATTGTCAGAACCATTTGGAG-3’
TfR-HPC4	III	5’-CGGTACCACCATGAGGCTGGCCGTGGGCGCCCTG-3’	IV	5’-GTCTAGATTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCAAACTCATTGTCAATGTCCCAA-3’
Cosmc with N-TfR	III	5’-CGGTACCACCATGAGGCTGGCCGTGGGCGCCCTG-3’	II	5’-GCTCTAGACTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCGTCATTGTCAGAACCATTTGGAG-3’
	V	5’-GCTGACAATAACACAAAGGCCAATGTCACAAAACCAAAAAGGGGTGTGATGCTTGGAAGCATTTTC-3’	VI	5’-ATAGTGATCAAAGCACAGAAAATGCTTCCAAGCATCACACCCCTTTTTGGTTTTGTGACATTGGC-3’

TfR with N-Cosmc	VII	5’-CGGATCCACCATGCTTTCTGAAAGCAGCTCCTTTTTGAAGTGTAGTGGAAGTATCTGCTATGGG-3’	IV	5’-GTCTAGATTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCAAACTCATTGTCAATGTCCCAA-3’
TfR with C-Cosmc	III	5’-CGGTACCACCATGAGGCTGGCCGTGGGCGCCCTG-3’	II	5’-GCTCTAGACTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCGTCATTGTCAGAACCATTTGGAG-3’
	VIII	5’-TTTTTCTTGATTGGATTTATGATTGGCTACTTGGGCTATTGTAGGATTGGTCATGGAAATAGAATG-3’	IX	5’-ATGATGCTCATGGTGGTGCATTCTATTTCCATGACCAATCCTACAATAGCCCAAGTAGCCAATCAT-3’
Cosmc with C-TfR	I	5’-CAATGAAAAGACTTAACAG CCT-3’	IV	5’-GTCTAGATTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCAAACTCATTGTCAATGTCCCAA-3’
	X	5’-GGAAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAAAGGGGTAGAACCAAAAACTGAG-3’	XI	5’-TCCTGCCAGTCTCTCACACTCAGTTTTTGGTTCTACCCCTTTAATGTGTCCTAGCATAGTGATCAA-3’
TfR with T-Cosmc	III	5’-CGGTACCACCATGAGGCTGGCCGTGGGCGCCCTG-3’	IV	5’-GTCTAGATTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCAAACTCATTGTCAATGTCCCAA-3’
	V	5’-GCTGACAATAACACAAAGGCCAATGTCACAAAACCAAAAAGGGGTGTGATGCTTGGAAGCATTTTC-3’	VI	5’-CATAGTGATCAAAGCACAGAAAATGCTTCCAAGCATCACACCCCTTTTTGGTTTTGTGACATTGGC-3’
Cosmc with T-TfR	VII	5’-GCTGACAATAACACAAAGGCCAATGTCACAAAACCAAAAAGGGGTGTGATGCTTGGAAGCATTTTC-3’	II	5’-GCTCTAGACTACTTGCCGTCGATCAGCCTGGGGTCCACCTGGTCCTCGTCATTGTCAGAACCATTTGGAG-3’

The six chimeric constructs were prepared using the following strategy. Plasmids A and B are color-coded and shown as two paired strands. Synthetic oligonucleotide primers are also color-coded and shown as single strands, with half arrowheads indicating the direction(88). The intermediate PCR-amplified products C and D are also shown as

two paired strands and color-coded according to the plasmid template and primers. They are derived from plasmid A and B, respectively (88).

In the first PCR, Plasmid A was amplified by high-fidelity PCR employing *pfu* polymerase and synthetic oligonucleotide primers a and b (88). Plasmid B was amplified by high-fidelity PCR employing *pfu* polymerase and synthetic oligonucleotide primers c and d. The conditions for the high−fidelity PCR amplification were: 1 cycle at 98°C for 30s, 35 cycles of PCR at 98°C for 10 s, 62°C for 30 s, and 72°C for 3−4 min (30s per kilobase DNA fragment to be amplified) in a volume of 50 μl with 1μl plasmid DNA template, 1 U *Pfu* polymerase, 10 pmol of each primer, 5 μl mixed dNTP at 2.5 mM concentration, and 10 μl 5x *Pfu* buffer (containing 1.0 mM Mg^{2+}), and 1 cycle at 72°C for 5 min. Then, the products from each individual PCR were analyzed on agarose gels and purified, mixed in an asymmetric ratio for a second PCR; the C strand at its 3'end and the D strand at its 5' ends have their overlap region and could pair with each other. Extension of this overlap by DNA polymerase created the full−length chimeric molecule E (88). The new strand E then acted as a template to make the final PCR product F by high-fidelity PCR employing *pfu* polymerase and synthetic oligonucleotide primers a and d (88). The optimal conditions for the second PCR were: 1 μl of each product from the first PCR by the high−fidelity PCR amplification was mixed with 10 pmol of synthetic oligonucleotide primers. The mixture was subjected a PCR in a volume of 50 μl containing 1 U of *pfu* polymerase, 5 μl mixed dNTP at 2.5 mM concentration, and 10 μl 5x pfu buffer (containing 1.0 mM Mg^{2+})(88). The PCR cycling parameters were the same as in the first PCR.

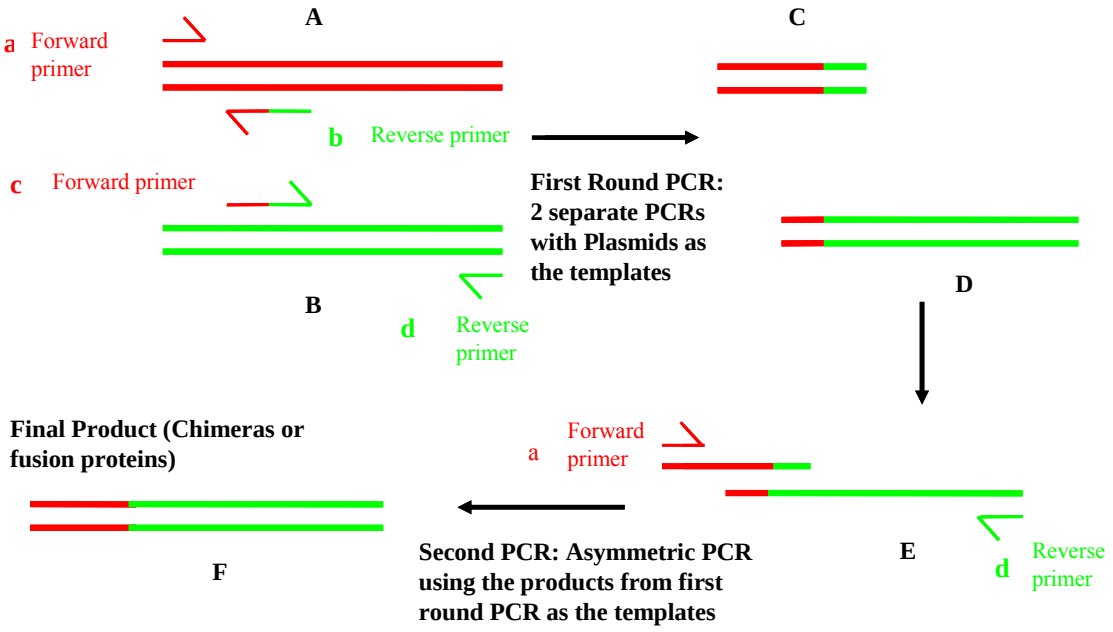


Figure 5. **Schematic diagram of the PCR-mediated recombination strategy used in making chimeric constructs.**

2.2.3 Site-Directed Mutagenesis

Wild type Cosmc and the chimera with transmembrane domain from Cosmc and the rest from TfR are used as templates to make mutations either from cysteine to alanine or from cysteine to serine using QuickChange™ site-directed mutagenesis kit following Manufacture's protocol. For Cys-to-Ala site directed mutagenesis, we used forward primer: 5'GGAAGCATTTTCGCTGCTTTGATC-3' and reverse primer: 5'GATCAAAGCAGCGAAAATGCTTCC-3'. For Cys-to-Ser site directed mutagenesis, we sued forward primer: 5'GGAAGCATTTTCAGTGCTTTGATC-3' and reverse primer:5'GATCAAAGCACTGAAAATGCTTCC-3'. The recombinant plasmids containing the desired mutations were confirmed by DNA sequencing.

2.2.4 Cell culture and transfection

Monolayer COS7 cells were cultured at 37°C in 5% CO_2 and in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with10% fetal bovine serum (FBS), penicillin, and streptomycin at 100 µg/ml. One day prior to transfection, the cells were seeded into a 10-cm dish and cultured in the complete media for overnight to reach to 50-60% confluency. Then cells were transfected with Fugene6 transfection reagent (Roche), according to the manufacturer's instructions.

2.2.5 Immunofluorescent staining of COS7 cells

COS7 cells were cultured on a chambered slide and transiently transfected with the expression constructs using Fugene 6 transfection reagent according to the manufacturer's protocol. At 48 h post transfection, cells were washed with TBS and fixed with 4% PFA on ice for 1 h and permeabilized with 0.1% Triton X-100 for 45 min on ice. After blocking with 1% BSA in TBS for 1 h at RT, the cells were incubated with primary antibodies for 1 h at RT. The cells were washed with TBS three times and incubated with Alexa Fluor–labeled secondary antibodies at RT for 1 h. Cells were then washed four times with TBS and mounted with Prolong Antifade Media (Invitrogen). After drying at RT for 12–16 h, cells were visualized on a confocal microscope (TCS NT; Leica) at RT under 40x Plan Fluotar 1.0 NA oil immersion or 100x Plan APO 1.4 NA oil immersion objective lenses. The images were maximum projection collected with a pinhole of 1 using 0.5-µm step size. Images were analyzed using the TCS and Volocity software (Leica).

2.2.6 Subcellular fractionation

COS7 cells grown to 80% confluence in 10 cm^2 dishes and transiently transfected with the expression constructs for 48 h were harvested and washed with cold PBS. Cells were homogenized in 25 mM HEPES, pH 7.5 containing 250 mM sucrose. Then post-nuclear supernatants (PNS) were made by centrifugation at 20,000 *g* for 30 min. The concentration of sucrose in PNS was adjusted to 40% (wt/vol) and loaded on 60, 50, 30, and 20% sucrose gradient. After centrifugation at 100,000 *g* for 20 h, 18 fractions were collected from the bottom of the tube. The fractions were analyzed by Western blotting with antibodies indicated in the results.

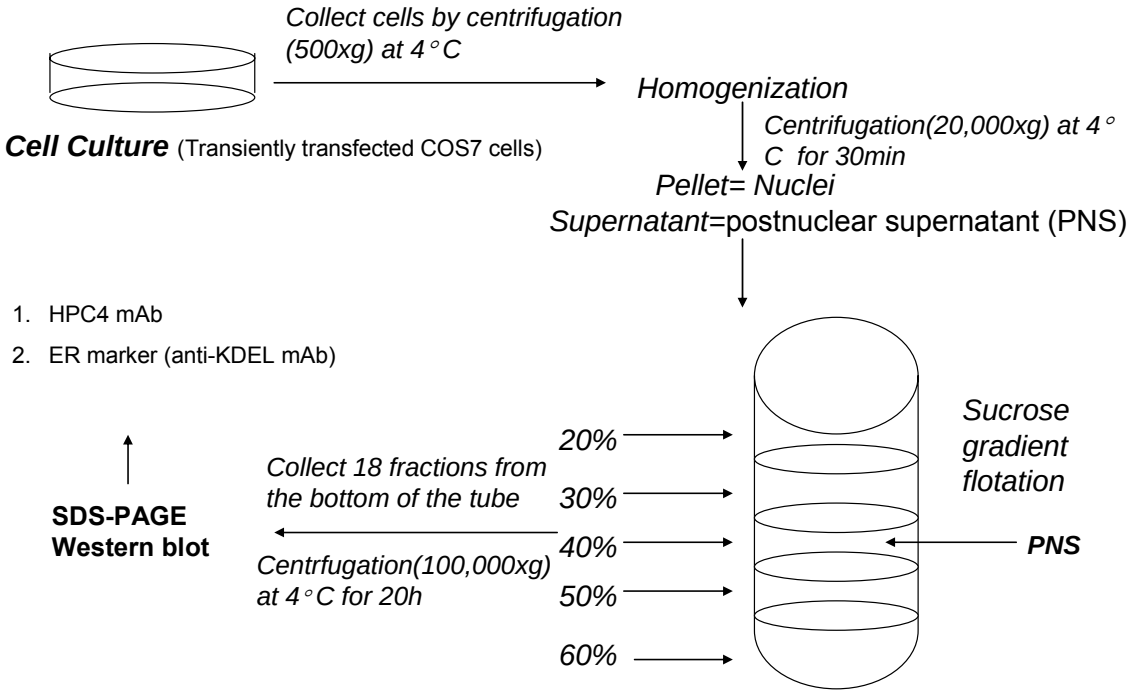


Figure 6. Schematic diagram of the subcellular fractionation.

2.2.7 Preparation of cell extracts

Transfected cells were resuspended in an appropriate volume of TBS buffer including 1mM $CaCl_2$ and proteinase inhibitor mixture (Roche Molecular Biochemicals) and homogenized by sonication on an ice-bath for 5 sec four times. The PNS were obtained by centrifugation of homogenate at 700 × g for 10 min, and the extracts were obtained by adding 1% Triton X-100 to the supernatant and solubilizing on ice for 30 min.

2.2.8 Western Blot

The protein samples were first separated on 4-20% SDS-PAGE, and then transferred onto a nitrocellulose membrane by electro-blotting at 30 V for 1.5 h using a Bio-Rad Mini Trans-Blot apparatus. The membrane was blocked with 5% low-fat milk. Primary antibody (10 μg/ml HPC4 mAb) was added onto the membrane for 1 h at room temperature. For HPC4 blotting, the membrane was washed 5 times with 20 mM Tris-HCl, 300 mM NaCl, and 1 mM $CaCl_2$, pH 7.2. After washing, the membranes were incubated with 4 ml of 0.17 ng/ml HRP-conjugated goat anti–mouse IgG at 23°C for 45 min. Membranes were washed 5 times with 20 mM Tris-HCl, 300 mM NaCl, and 1 mM $CaCl_2$, pH 7.2, and incubated with 3 ml of SuperSignal West Pico Chemiluminescent Substrate at RT for 1 min. The blot was exposed to film (Denville Scientific, Inc.) for 1 min, and developed by autoradiography. The membranes were stripped by washing with 2% SDS and 1% 2-mercaptoethanol in deionized water for 30 min, followed by washing 3 times with 20 mM Tris-HCl and 300 mM NaCl, pH 7.2. The membranes were blocked with 1% milk and washed 3 times with 20 mM Tris-HCl and 300 mM NaCl, pH 7.2. Membranes were incubated with polyclonal mouse IgG against KDEL using 4 ml of 0.2

μg/ml antibody for 1 h. Membranes were washed 5 times with 20 mM Tris-HCl and 300 mM NaCl, pH 7.2, and incubated with 4 ml of 0.17 ng/ml HRP-conjugated goat anti–chicken IgG at 23°C for 45 min. Membranes were washed with 20 mM Tris-HCl and 300 mM NaCl, pH 7.2, and developed accordingly.

2.2.9 Proteasome inhibitor treatment

About 10^6 COS7 cells were seeded in T75 flasks and transiently transfected with the expression constructs using Fugene 6 transfection reagent according to the manufacturer's protocol. At 48 h after transfection, the cells were split into two plates and treated with 10 μM MG-132 (dissolved in 100% DMSO at 2 mM stock) or 0.5% DMSO in complete media for 12–14 h, respectively. The cells were harvested for Western blot.

2.2.10 *In vitro* Cosmc immunoprecipitation

Proteins were incubated with beads overnight while rotating at 4°C. The beads were collected using bench centrifugation and the beads were washed 5 times with 400 μl TBS washing buffer. The bead-bound material was eluted 5 times with 20 μl of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, 0.1% Triton X-100).

2.3 Results

2.3.1 Generation of Cosmc-TfR chimeras

To exam the role of the cytoplasmic, the TMD, and the luminal domains of Cosmc in the secretory pathway, a series of chimeric cDNAs were constructed by recombinant DNA techniques. Cosmc was fused to the equivalent region from a cell surface membrane

protein, transferrin receptor (TfR). TfR is a type II transmembrane protein found in all animal cells. It was chosen as a reporter because, like Cosmc, it has a short cytoplasmic domain, a single membrane-spanning domain, and a long extracellular domain.

Six constructs for expression of chimeras have been made. Chimera 1 contains the TfR cytoplasmic domain and Cosmc TMD and (lumenal domain) LD (#1- TfR/Cosmc/Cosmc). Chimera 2 contains the Cosmc cytoplasmic domain and TfR TMD and extracellular domain (ECD) (#2- Cosmc/TfR/TfR). Chimera 3 contains the Cosmc luminal domain and the TfR cytoplasmic and TMD (#3- TfR/TfR/Cosmc). Chimera 4 contains the TfR luminal domain and the Cosmc cytoplasmic and TMD (#4- Cosmc/Cosmc/TfR). Chimera 5 contains the Cosmc TMD and the TfR cytoplasmic and luminal domain (#5- TfR/Cosmc/TfR). Finally, chimera 6 contains the TfR transmembrane domain and the Cosmc cytoplasmic and luminal domain (#6- Cosmc/TfR/Cosmc). Each construct has a luminal-terminal HPC4 tag. Figure 7 displays each construct.

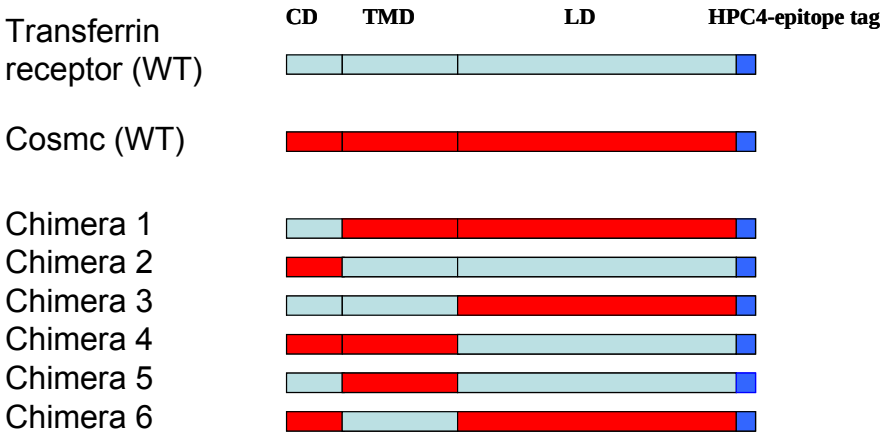


Figure 7. **Schematic diagram of the starting proteins and the generated various chimeras used in this study. The major protein domains, CD (cytoplasmic domain); TMD (transmembrane domain); LD (luminal domain); and HPC4 epitope tag in blue, are represented by closed boxes.**

2.3.2 Wild type Cosmc localizes in the ER

COS7 cells were transiently transfected with full-length, HPC4-epitope tagged wild type Cosmc (wtCosmc), and stained with HPC4 mAb by immunofluorescence. As shown in Figure 8A-C, Cosmc (red) was expressed in a perinuclear pattern in the transfected cells, with a very similar pattern to the ER marker Calnexin (green), which stained all cells. A merge of the staining image gave a yellow color in the Cosmc expressing cells, indicating that Cosmc and Calnexin are co-localized. This result is consistent with our previous data, demonstrating that wtCosmc mainly localizes in ER(56).

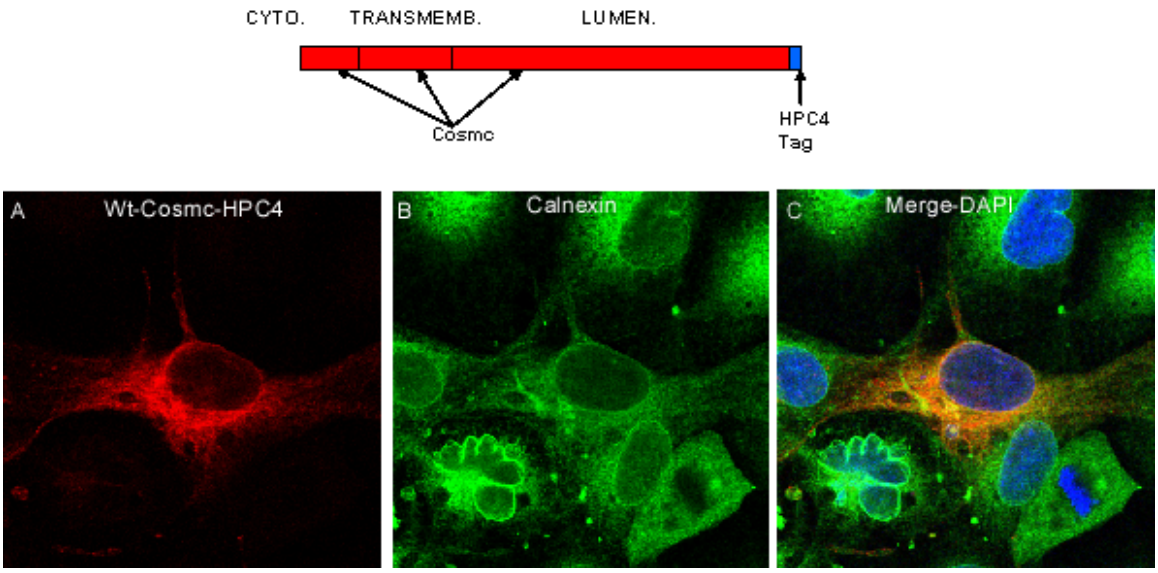


Figure 8. **Localization of the wtCosmc. (A-C) Immunofluorescent staining of the wtCosmc. COS-7 cells grown on chambered slides were transiently transfected with C-terminally HPC4-tagged wtCosmc. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody, rabbit anti-Calnexin (green)**

2.3.3 Wild type TfR localizes in the plasma membrane

COS7 cells were transiently transfected with plasmid expressing the wild type TfR (wtTfR), and immofluorescently stained with anti-HPC4 and anti-Calnexin antibody in the presence and absence of detergent Triton X-100 for permeabilization.

The wtTfR (red) was observed as bright plasma membrane staining pattern under both permeable and non-permeable conditions, while Calnexin (green) was apparent only when cells were permeablized by Triton X-100, indicating its intracellular localization. A minor staining in the ER also observed with for wtTfR, which could be due to the high expression this protein in Cos-7 cells or overall fluorescence in many membranes including endosomes (Fig. 9A-D). To further confirm the localization of the wtTfR, we performed subcellular fractionation on sucrose gradients of COS7 cells expressing wtTfR-HPC4. wtTfR-HPC4 was recovered primarily in fractions 5-15 with major bands in fractions 6~10, while GRP78/94 were mainly in fractions 8~12 as shown in the Western Blot (Fig. 9E). wtTfR is first synthesized in the ER and then sorted to the plasma membrane through the secretory pathway. Thus, it is not surprising to see some portion of wtTfR in the ER as well as plasma membranes. Overall, the IF staining and subcellular fractionation results are consistent each other and indicate that wtTfR is mainly on plasma membrane of COS-7 cells, as expected.

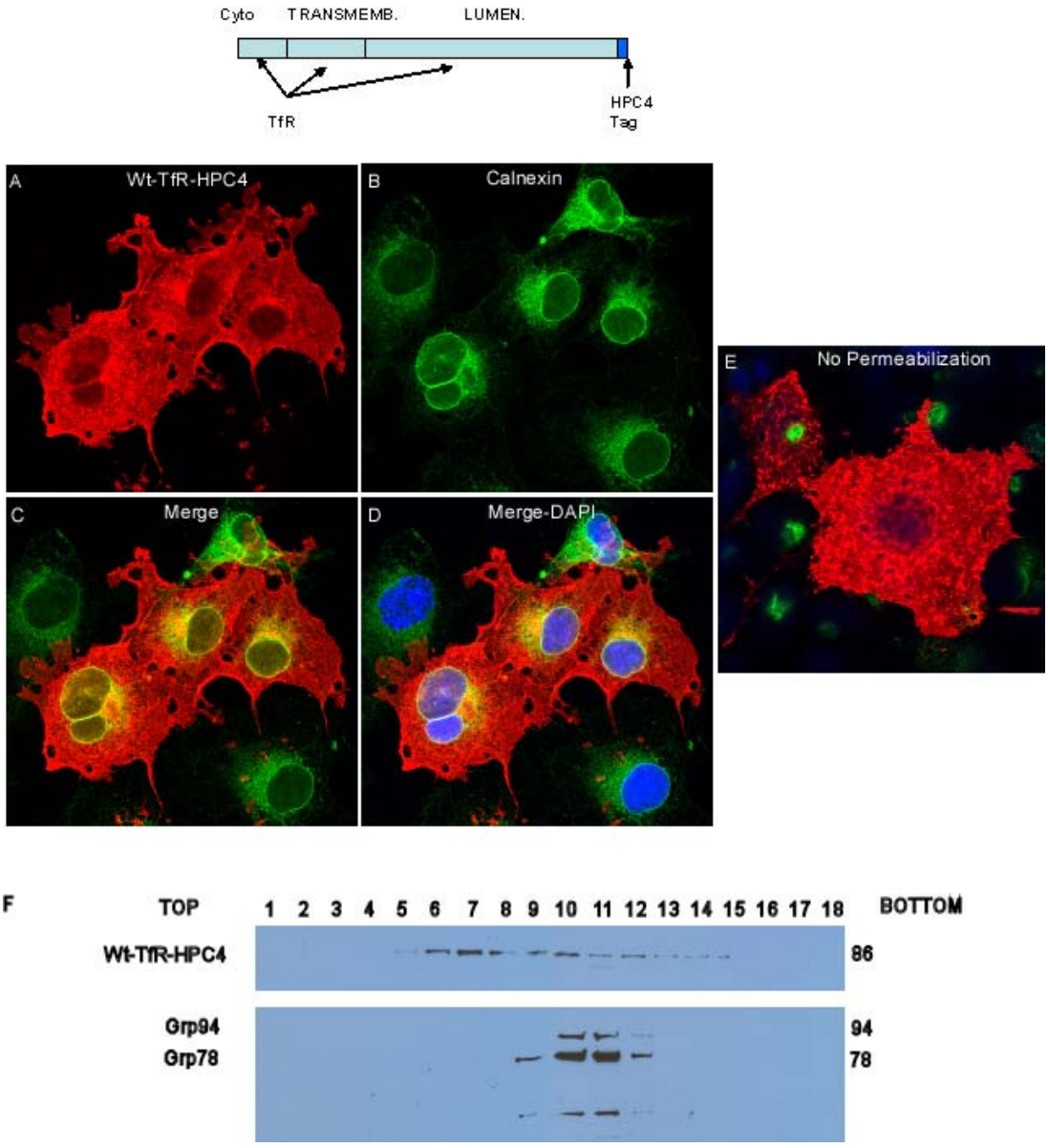


Figure 9. **Localization of the wtTfR. (A-D) Immunofluorescent staining of the wtTfR. COS-7 cells grown on chambered slides were transiently transfected with C-terminally HPC4-tagged wtTfR TfR. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody, rabbit anti-Calnexin (green). (E)The transfected cells were not treated with Triton-100 to permeabilize the cell membrane, but still stained with mouse anti-HPC4 (red) antibody and rabbit anti-Calnexin (green). (F) Sucrose gradient subcellular fractionation. COS-7 cells transiently transfected with HPC4-tagged wild type TfR were harvested, homogenized and removed the nuclei. The postnuclear supernant (PNS) was applied to a sucrose gradient and 18 fractions were obtained after ultracentrifugation. Proteins from each fraction were analyzed on Western Blot with anti-HPC4 and anti-KDEL antibodies.**

2.3.4 Cosmc/TfR/TfR (Construct #2) localizes in the plasma membrane

COS7 cells were transiently transfected with the construct expressing chimera Cosmc/TfR/TfR (Construct #2)**,** and then the cellular localization was examined by immunofluorescence with anti-HPC4 antibody as shown in Figure 10A-D. The ER marker was visualized by antibody to an ER-resident protein Calnexin.

Cosmc/TfR/TfR (Construct #2) (red) almost equally displayed on a whole cell with a clear edges of the plasma membrane , as seen in the wild type TfR staining, it was also observed as perinuclear and punctuate pattern in some cells. These results indicate that this chimera was mainly localized on the plasma membrane, and partially retained in the ER and Golgi apparatus. The merged image with Calnexin (green) confirmed the partial co-localization of Cosmc/TfR/TfR with the ER marker, and the partial ER-localization of this fusion protein. To further confirm its plasma membrane localization, the transfected COS-7 cells were also stained with anti-HPC4 at non-permeablization condition. We observed that the Cosmc/TfR/TfR (Construct #2) was similar stained on the whole cell with a pattern similar to observed in the permeabilized cells, besides the perinuclear staining (Fig. 10E). Another line of evidence showing that Cosmc/TfR/TfR was mainly plasma membrane-localized was obtained from subcellular fractionation/Western Blotting experiment. As shown Figure 10F, Cosmc/TfR/TfR-HPC4 from transfected COS-7 cells was recovered primarily in fractions 4-13, where GRP78/94, the ER proteins were mainly in fractions 10-12, which is similar to the wild type TfR. All data demonstrated chimera #2 was mainly localized on plasma membrane. These results indicated that the cytoplasmic domain of Cosmc is not sufficient to retain

the TfR in the ER, suggesting that the cytoplasmaic domain of Cosmc is not the main ER-localization determinants for Cosmc.

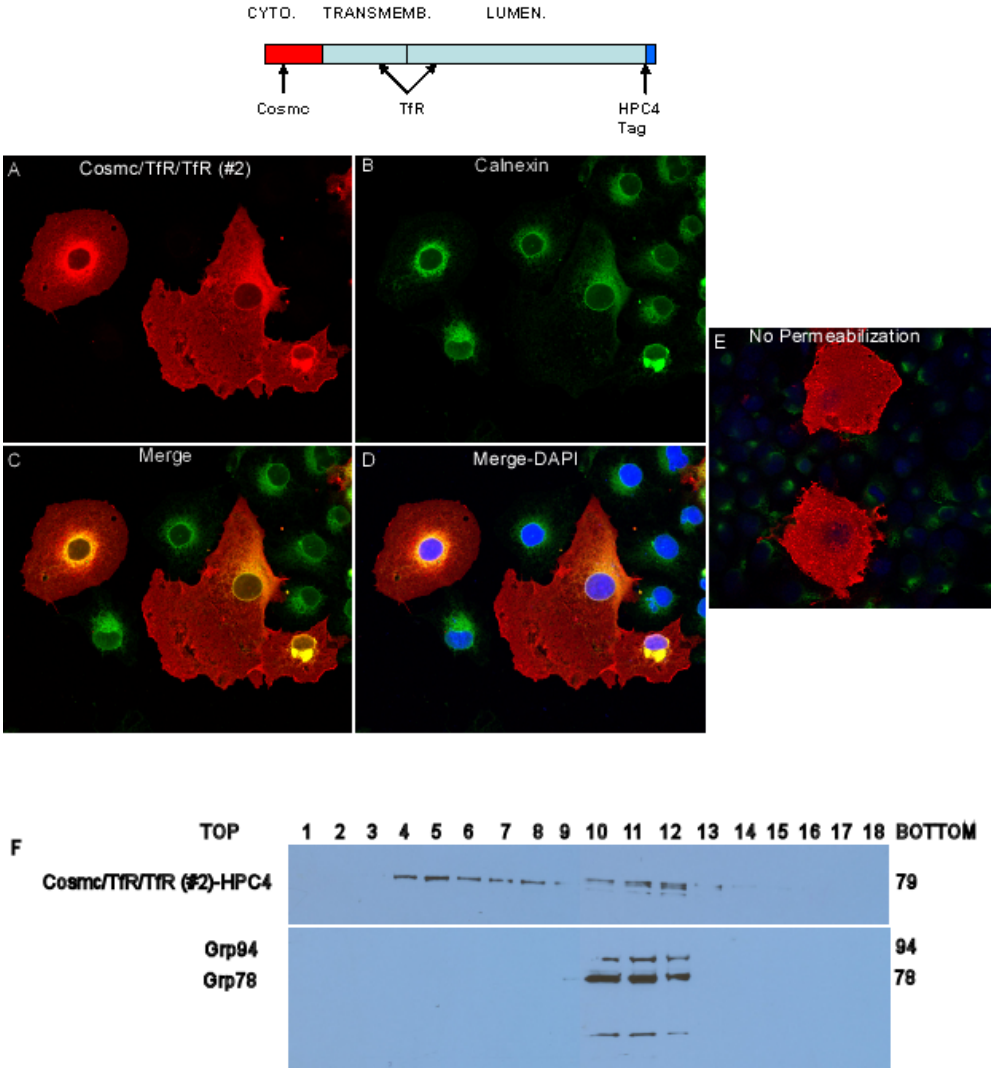


Figure 10. **Localization of the construct Cosmc/TfR/TfR. (A-D) Immunofluorescent staining of Cosmc/TfR/TfR. COS-7 cells grown on chambered slides were transiently transfected with Cosmc/TfR/. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody, rabbit anti-Calnexin (green). (E)The transfected cells were not treated with Triton-100 to permeabilize the cell membrane, but still stained with mouse anti-HPC4 (red) antibody and rabbit anti-Calnexin (green). (F) Sucrose gradient subcellular fractionation. COS-7 cells transiently transfected with HPC4-tagged Cosmc/TfR/TfR were harvested, homogenized and removed the nuclei. The postnuclear supernant (PNS) was applied to a sucrose gradient and 18 fractions were obtained after ultracentrifugation. Proteins from each fraction were analyzed on Western Blot with anti-HPC4 and anti-KDEL antibodies.**

2.3.5 TfR/TfR/Cosmc (Construct #3) localizes in the Golgi

To test whether the luminal domain of Cosmc is important in ER-localization, we made a construct expressing the chimera protein TfR/TfR/Cosmc (Construct #3). COS7 cells were transiently transfected with the construct, and then the cellular localization was examined by immunofluorescence with the HPC4 antibody as shown in Figure 11A-D. The Golgi marker was visualized by antibody to a Golgi-resident protein Giantin.

Surprisingly, TfR/TfR/Cosmc (Construct #3) (red) revealed a punctate pattern in many cells that was coincident with the localization of Giantin (green), a Golgi marker, which stained in every cell. Merge of the staining image gave a yellow color in every TfR/TfR/Cosmc (Construct #3) stained cells, indicating the co-localization of those two proteins. To further establish the localization of TfR/TfR/Cosmc (Construct #3), we performed subcellular fractionation on sucrose gradients of COS7 cells expressing TfR/TfR/Cosmc-HPC4 (Construct #3). TfR/TfR/Cosmc (Construct #3) was present in fractions 8-9 as detected by anti-HPC on Western blot (Fig.11E), which is different from fractions 10-12 containing the ER marker Grp78/94. This result is consistent with the immunofluorescent staining (IF) data further confirming the Golgi-localization of construct #3. These results indicated that the luminal domain of Cosmc is not sufficient to retain the TfR in the ER, suggesting that the luminal domain of Cosmc is not the main ER-localization determinant for Cosmc. However, this luminal domain has the ability to retain Construct #3 in the Golgi apparatus.

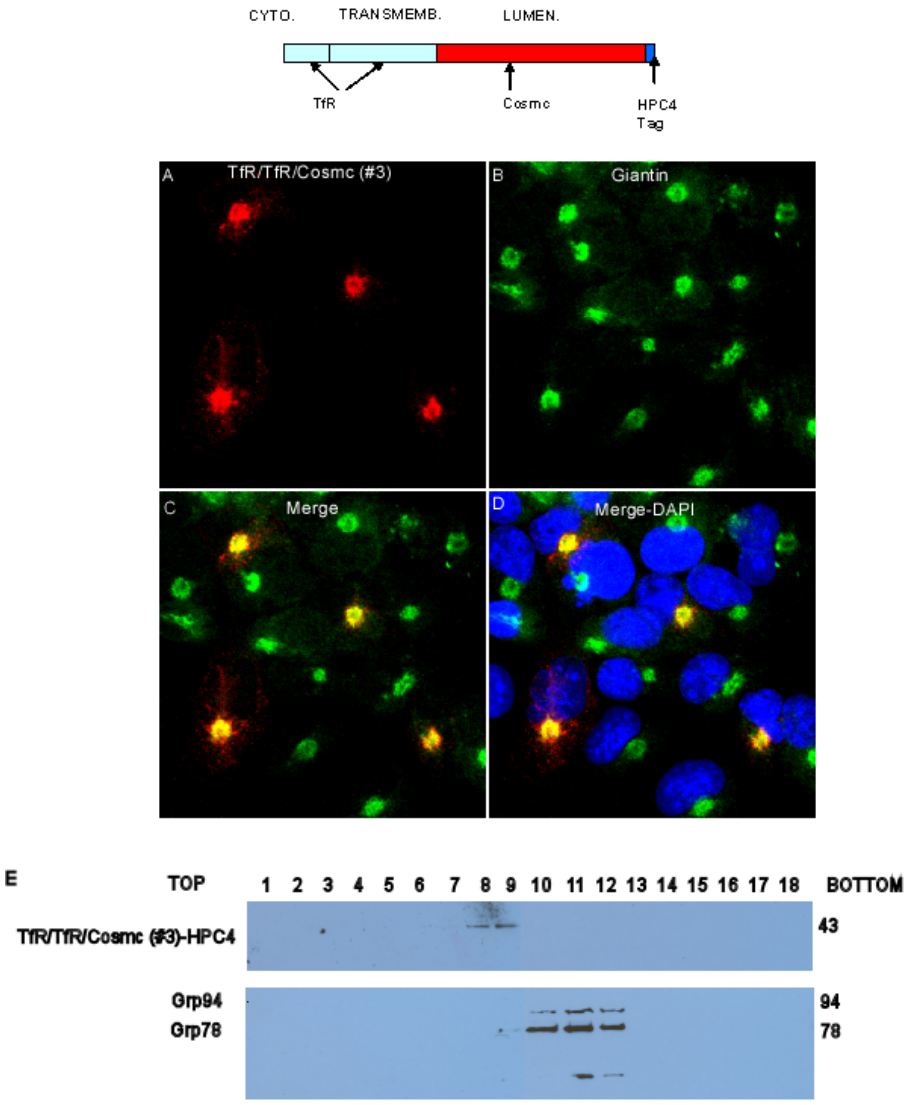


Figure 11. **Localization of the Construct TfR/TfR/Cosmc. (A-D) Immunofluorescent staining of TfR/TfR/Cosmc. COS-7 cells grown on chambered slides were transiently transfected with TfR/TfR/Cosmc. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody and rabbit anti-Giantin (green) antibodies. (E) Sucrose gradient subcellular fractionation. COS-7 cells transiently transfected with HPC4-tagged TfR/TfR/Cosmc were harvested, homogenized and removed the nuclei. The postnuclear supernant (PNS) was applied to a sucrose gradient and 18 fractions were obtained after ultracentrifugation. Proteins from each fraction were analyzed on Western Blot with anti-HPC4 and anti-KDEL antibodies.**

2.3.6 Cosmc/Cosmc/TfR (Construct #4) localizes in the ER

COS7 cells were transiently transfected with the construct expressing Cosmc/Cosmc/TfR(Construct #4), and then the cellular localization was examined by immunofluorescence with the HPC4 antibody as shown in Figure 12A-D. The ER marker was visualized by antibody to an ER-resident protein Calnexin.

Cosmc/Cosmc/TfR (Construct #4) (red) was observed in a perinuclear pattern in some cells, similar to staining with ER marker Calnexin (green) which stained in every cell. Merge of the staining image gave a yellow color in the Cosmc/Cosmc/TfR stained cells, indicating the co-localization of those two proteins. To further establish the localization of Cosmc/Cosmc/TfR (Construct #4), we performed subcellular fractionation on sucrose gradients of COS7 cells expressing Cosmc/Cosmc/TfR-HPC4 (Construct #4). Cosmc/Cosmc/TfR-HPC4 (Construct #4) was recovered primarily in fractions 9-12 as detected by anti-HPC on Western blot, corresponding to fractions 9-12 containing the ER markers GRP78 and GRP94 (Fig.12E). This result is consistent with the immunofluorescent staining (IF) data further confirming the ER-localization of chimera #4. These results indicated that the cytoplasmic and transmembrane domains of Cosmc are sufficient to retain the TfR in the ER.

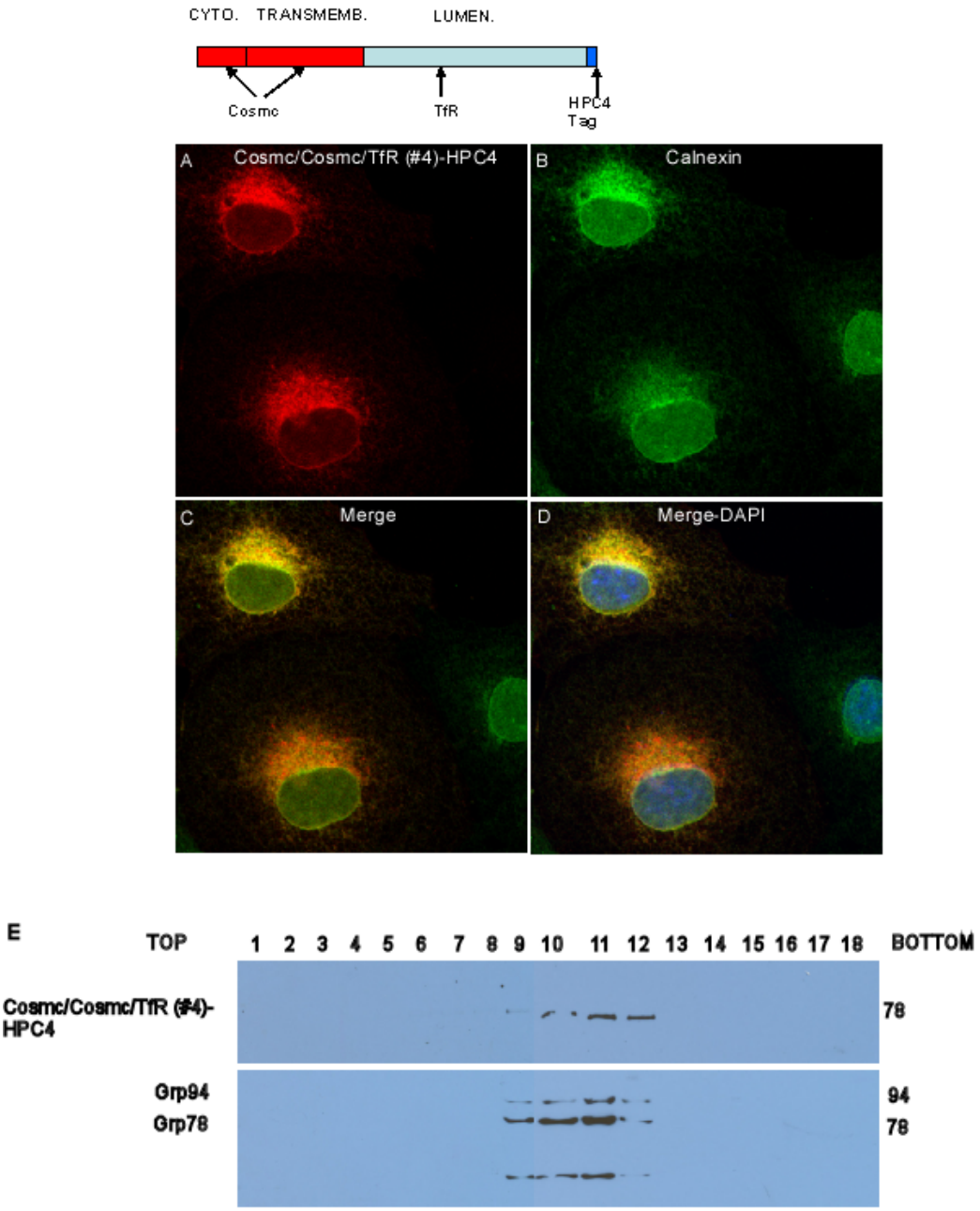


Figure 12. **Localization of the construct Cosmc/cosmc/TfR. (A-D) Immunofluorescent staining of Cosmc/Cosmc/TfR fusion proteins. COS-7 cells grown on chambered slides were transiently transfected with Cosmc/Cosmc/TfR. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody and rabbit anti-Calnexin (green). (E) Sucrose gradient subcellular fractionation. COS-7 cells transiently transfected with HPC4-tagged Cosmc/Cosmc/TfR were harvested, homogenized and removed the nuclei. The postnuclear supernant (PNS) was applied to a sucrose gradient and 18 fractions were obtained after ultracentrifugation. Proteins from each fraction were analyzed on Western Blot with anti-HPC4 and anti-KDEL antibodies.**

2.3.7 TfR/Cosmc/TfR (Construct #5) localizes in the ER

To test whether the TMD of Cosmc is the ER-localization determent for Cosmc or not, we made a construct expressing the chimera protein of TfR/Cosmc/TfR (Construct #5). COS7 cells were transiently transfected with the construct, and then the cellular localization was examined by immunofluorescence with the HPC4 antibody as shown in Figure 13A-D. The ER marker was visualized by antibody to an ER-resident protein Calnexin.

TfR/Cosmc/TfR (Construct #5) (red) was observed in a perinuclear pattern in some cells, similar to staining with ER marker Calnexin (green) which stained in every cell. Merge of the staining image gave a yellow color in the TfR/Cosmc/TfR stained cells, indicating the co-localization of those two proteins. To further establish the localization of TfR/Cosmc/TfR (Construct #5), we performed subcellular fractionation on sucrose gradients of COS7 cells expressing Chimera #5. TfR/Cosmc/TfR-HPC4 (Construct #5) was recovered primarily in fractions 10-12 as detected by anti-HPC on Western blot, corresponding to fractions 10-12 containing the ER markers GRP78 and GRP94 (Fig.13E). This result is consistent with the immunofluorescent staining (IF) data further confirming the ER-localization of chimera #5. These results indicate that the TMD alone of Cosmc is sufficient to retain the TfR in the ER, suggesting that the transmembrane domain of Cosmc is the main ER-localization determinant for Cosmc.

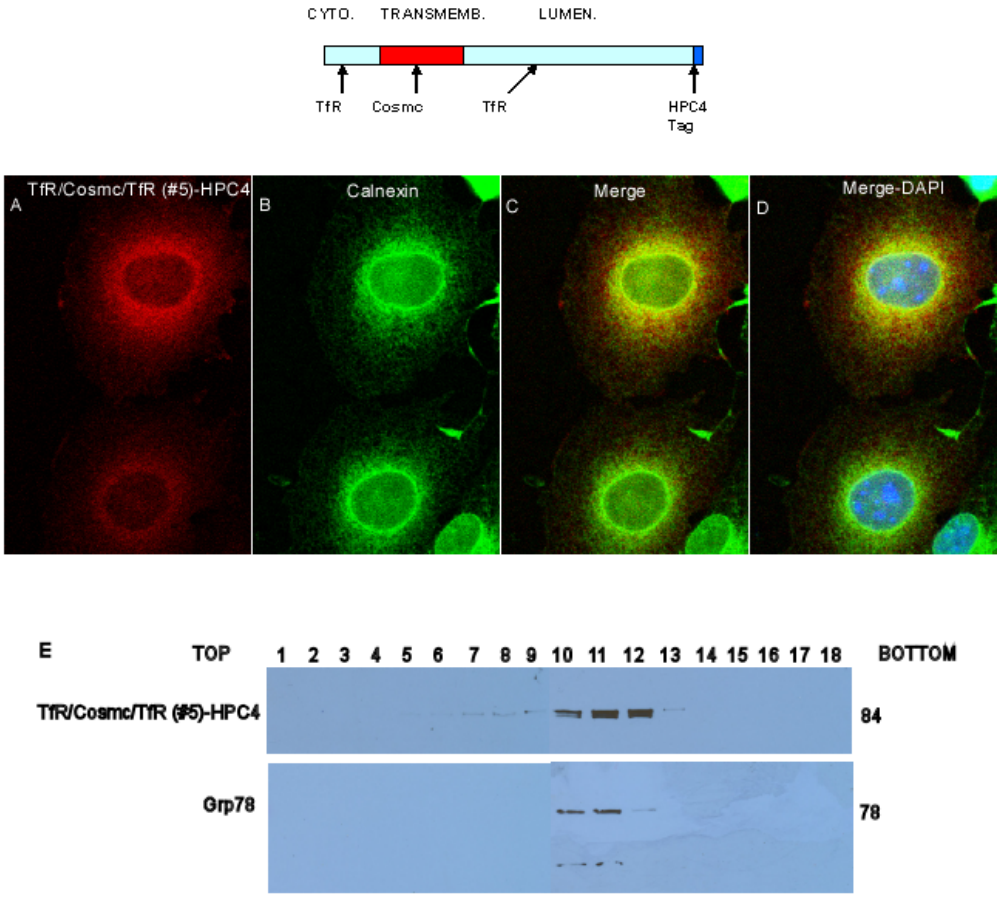


Figure 13. **Localization of the construct TfR/Cosmc/TfR. (A-D) Immunofluorescent staining of TfR/Cosmc/TfR. COS-7 cells grown on chambered slides were transiently transfected with TfR/Cosmc/TfR. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody and rabbit anti-Calnexin (green). (E) Sucrose gradient subcellular fractionation. COS-7 cells transiently transfected with HPC4-tagged TfR/Cosmc/TfR were harvested, homogenized and removed the nuclei. The postnuclear supernant (PNS) was applied to a sucrose gradient and 18 fractions were obtained after ultracentrifugation. Proteins from each fraction were analyzed on Western Blot with anti-HPC4 and anti-KDEL antibodies.**

2.3.8 Cosmc/TfR/Cosmc (Construct #6) mainly localizes in the Golgi, but partially in the ER

COS7 cells were transiently transfected with the construct expressing Cosmc/TfR/Cosmc (Construct #6), and then the cellular localization was examined by immunofluorescence

with anti-HPC4 antibody as shown in Figure 14A-H. The ER marker was visualized by antibody to an ER-resident protein Calnexin. The Golgi marker was visualized by antibody to a Golgi-resident protein Giantin.

Cosmc/TfR/Cosmc (Construct #6) (red) observed in both a perinuclear localization similar to staining with the ER marker Calnexin (green) (Fig 14A-D) and a punctate pattern that was coincident with the localization of Giantin, a Golgi marker (green) (Fig 14E-H). Merge of the staining image gave a yellow color in the Cosmc/TfR/Cosmc stained cells, indicating the co-localization of those two proteins. To further establish the localization of Cosmc/TfR/Cosmc (Construct #6), we performed subcellular fractionation on sucrose gradients of COS7 cells expressing chimera #6. Cosmc/TfR/Cosmc (Construct #6) was observed in fractions 4-12 as detected by anti-HPC on Western blot, which is different from fractions 9-12 containing the ER markers GRP78 and GRP94 (Fig.14I). These results are consistent with the immunofluorescent staining (IF) data and indicated that the absence of the TMD of Cosmc is unable to retain the TfR in the ER, suggesting that the transmembrane domain of Cosmc is the main ER-localization determinant for Cosmc.

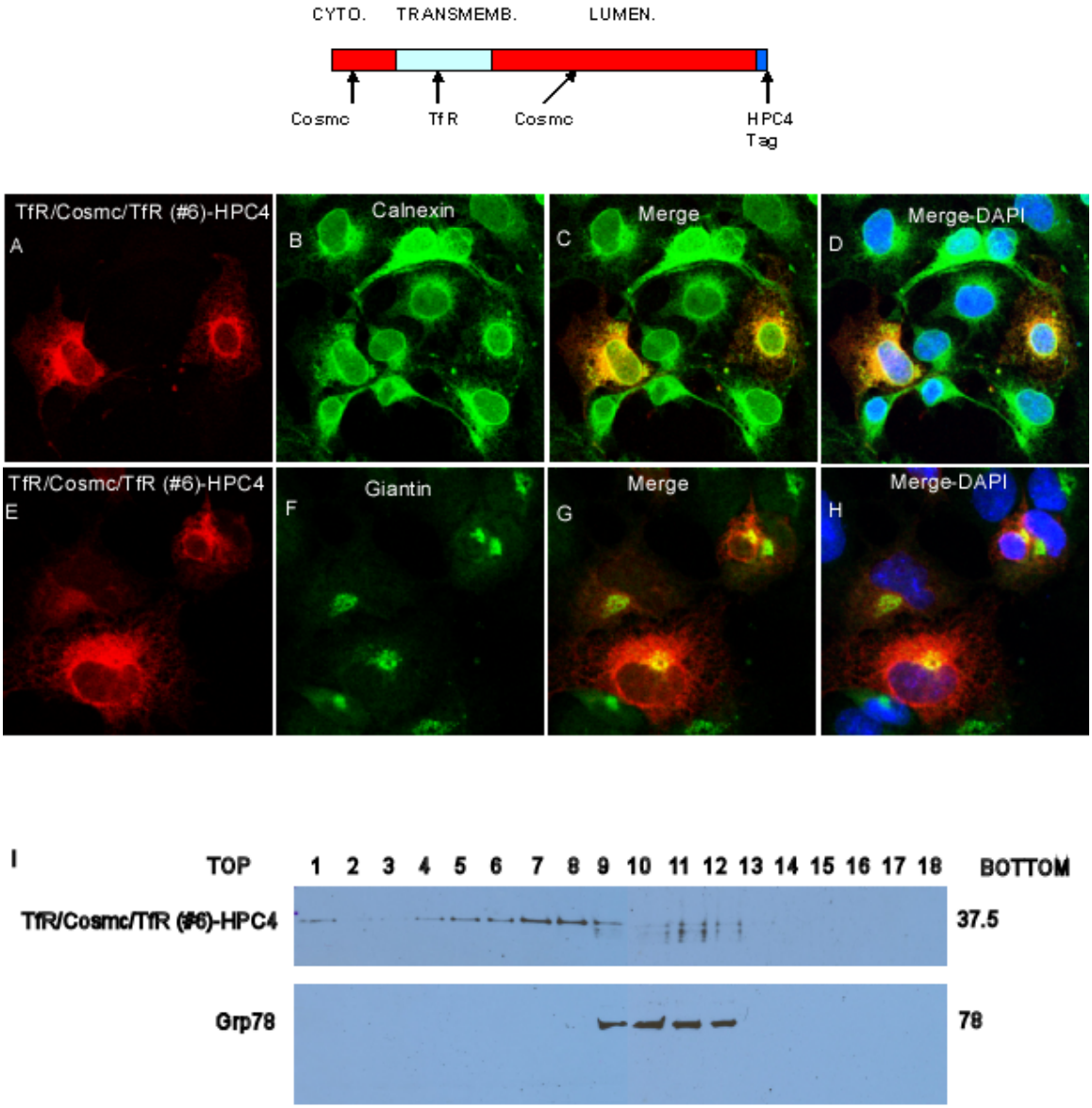


Figure 14. **Localization of the construct Cosmc/TfR/Cosmc. (A-H) Immunofluorescent staining of Cosmc/TfR/Cosmc. COS-7 cells grown on chambered slides were transiently transfected with Cosmc/TfR/Cosmc. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody, rabbit anti-Calnexin (green) or rabbit anti-Giantin (green) antibodies. (I)Sucrose gradient subcellular fractionation. COS-7 cells transiently transfected with HPC4-tagged Cosmc/TfR/cosmc were harvested, homogenized and removed the nuclei. The postnuclear supernant (PNS) was applied to a sucrose gradient and 18 fractions were obtained after ultracentrifugation. Proteins from each fraction were analyzed on Western Blot with anti-HPC4 and anti-KDEL antibodies.**

2.3.9 Construct TfR/Cosmc/Cosmc (#1) is unstable and degraded in proteosomes

COS-7 cells were transiently transfected with wtCosmc and TfR/Cosmc/Cosmc (Construct #1). However, unlike expression of other constructs, there was very little expression of the HPC4-tagged Construct #1 in any cell, thus suggesting that the recombinant protein might be degraded. To example this, after 48h transfection, both transfected COS-7 and non-transfected COS-7 cells were equally split into two plates. One was treated with proteasome inhibitor MG-132, while the other was treated overnight with dimethyl sulphoxide (DMSO) as a negative control. After collecting the cells, we examined expression of the Construct #1 using the mouse anti-HPC4 antibody in western blotting. There were no apparent bands present in non-transfected cells. Interestingly, there was very little expression of Construct #1 in the absence of MG-132, but the presence of the inhibitor significantly enhanced expression (Fig. 15). In addition, wtCosmc was clearly expressed in the absence of MG-132, but the presence of the inhibitor enhanced expression, indicating that some of the wtCosmc is also being degraded in cells by a proteasomal pathway. These data demonstrate that Construct #1 (fR/Cosmc/Cosmc) was degraded through the proteasome pathway and thus, it is not possible to examine its localization. Therefore, there are also no confocal images for TfR/Cosmc/Cosmc.

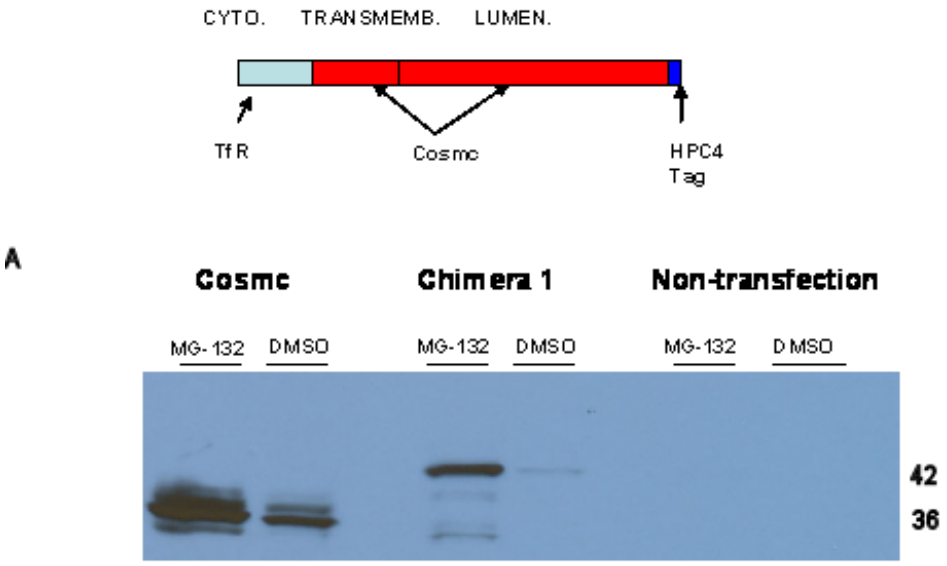
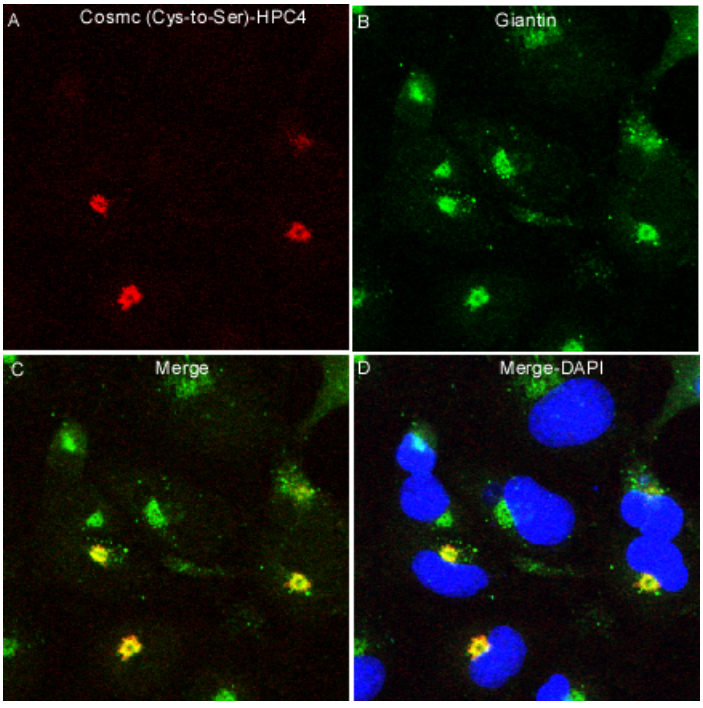


Figure 15. **Degradation of TfR/Cosmc/Cosmc (#1) through proteasome pathway. COS-7 cells transiently expressing wild type Cosmc and TfR/Cosmc/Cosmc were treated with 10μM MG-132 or DMSO for overnight and havested. The cell extracts were analyzed on SDS-PAGE by Western Blot with mouse anti-HPC4 antibody. MG-132 causes accumulation of both Cosmc and TfR/Cosmc/Cosmc, but DMSO only causes accumulation of Cosmc.**

2.3.10 The amino acid cysteine within the membrane spanning domain of Cosmc is required for the retention of the full-length Cosmc in the ER

The TMD of Cosmc does not contain any known ER retention motif. However, there is a single residue of cysteine within the TMD, which led us to explore whether this residue might contribute to the ER localization function of the TMD. To explore the potential contribution of this Cys residue in wtCosmc, the Cys was mutated to Ala or Ser by site-

directed mutagenesis. COS-7 cells were transfected with these two new constructs individually. Using immunofluorescence, the cellular localization of the constructs was assayed. Unexpectedly, the results from immunofluorescence showed a punctuate pattern of both two mutants (red), that was coincident with the localization of a Golgi marker, Giantin (green). These results demonstrate that the TMD of Cosmc is responsible for ER retention and that mutation of the single Cys residue in the TMD to either Ala or Ser causes the TMD to lose its ER localization function and the mutated proteins accumulate in the Golgi apparatus.



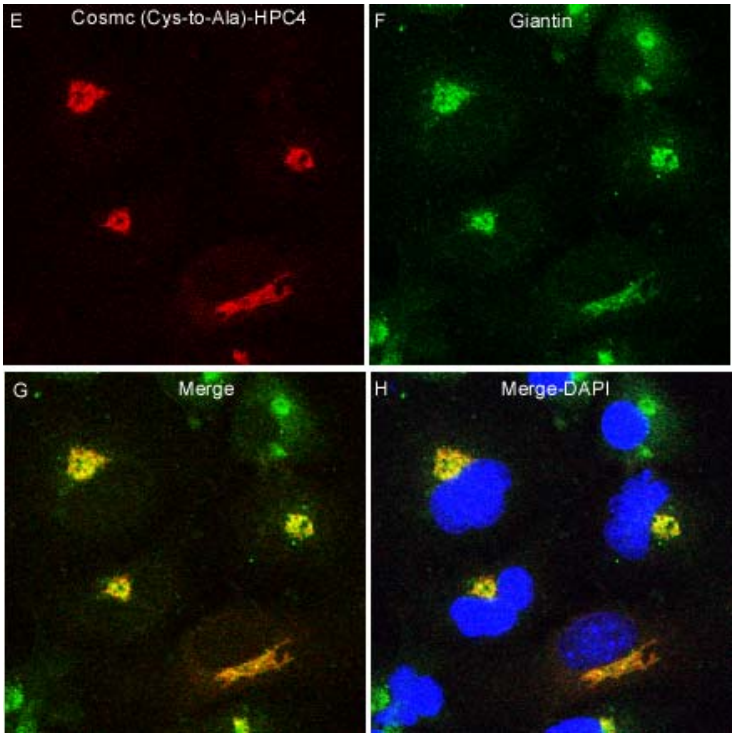


Figure 16. **Localization of the Cysteine mutant Cosmc. (A-D) Immunofluorescent staining of mutant Cosmc (Cys-to Ser). COS-7 cells grown on chambered slides were transiently transfected with Cosmc/TfR/Cosmc. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody and rabbit anti-Giantin (green) antibodies. (E-F) Immunofluorescent staining of mutant Cosmc (Cys-to-Ala). COS-7 cells grown on chambered slides were transiently transfected with Cosmc/TfR/Cosmc. After 24h, the post-transfected cells were then stained with mouse anti-HPC4 (red) antibody and anti-Giantin (green) antibodies.**

2.4 Discussion

The results presented in this study show that the TMD of Cosmc is sufficient to lead to the ER localization of a fusion protein based on the human transferrin receptor (TfR), a glycoprotein that normally localizes to the plasma membrane and endosomes. The chimera containing the Cosmc cytoplasmic domain (#2- Cosmc/TfR/TfR) largely accumulates on the cell surface, which means that there is no ER retrieval or retention signal in the cytoplasmic domain of Cosmc. The chimera containing the Cosmc luminal

domain (#3- TfR/TfR/Cosmc) was able to leave the ER and mainly localized in the Golgi, which suggests that the ER retention/retrieval signal is absent from the luminal domain of Cosmc, and that the cytoplasmic domain of Cosmc has some Golgi targeting ability. However, chimera #3 also did not localize to the cell membrane as did #2, therefore we speculate that there might be some accessory motif in the luminal domain for the ER localization of Cosmc. These two chimeras both lack the TMD of Cosmc, which can demonstrate the importance of TMD of Cosmc for its ER localization. The chimeras containing the Cosmc TMD (#4- Cosmc/Cosmc/TfR and #5- TfR/Cosmc/TfR) mainly accumulated in the ER, which confirms that the TMD of Cosmc is responsible for the retention/retrieval of Cosmc.

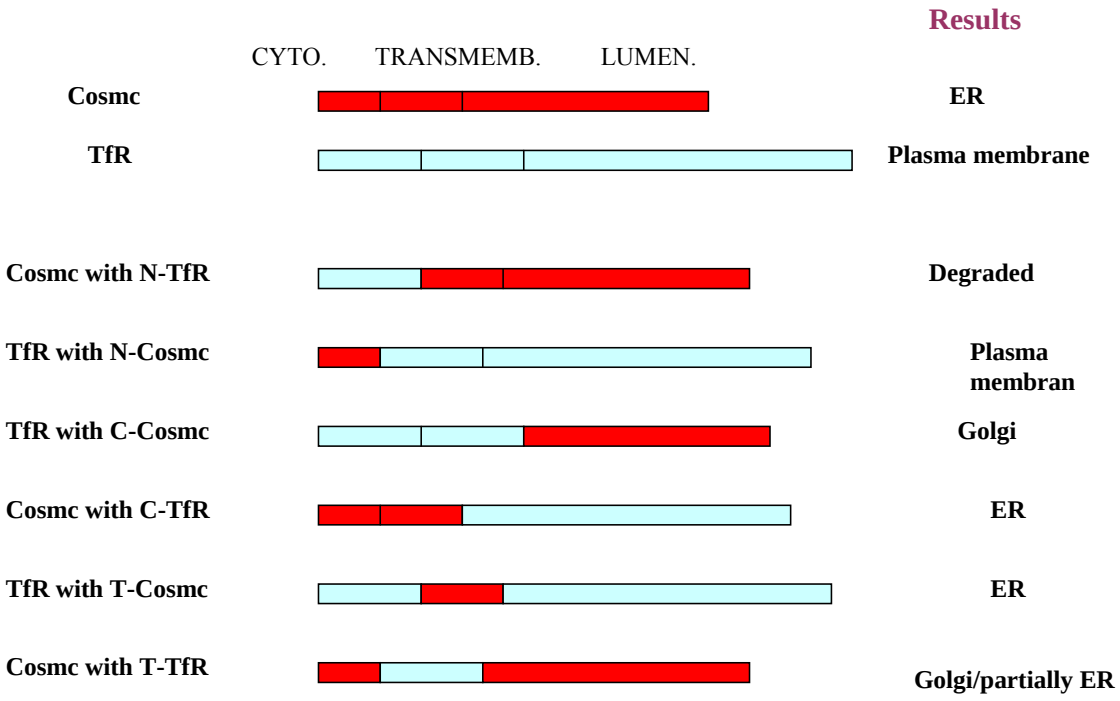


Figure 17. Summary of the results.

The molecular mechanisms for localization of many ER-resident proteins have been partly defined(64,70,72,81,89,90). However, the mechanisms for localization of proteins lacking such motifs are obscure. The genesis of our study is that Cosmc, an ER-localized chaperone, has no apparent motifs for ER retention and localization. The results of our study show that the TMD of Cosmc is critical to its ER localization, and that this TMD can confer ER localization to a non-ER protein, such as the human transferrin receptor.

Based on known ER and Golgi targeting motifs, we considered several possible models. In the first model, the length of the membrane-spanning domain of type II ER-resident proteins determines how a membrane protein is sorted to the lipid bilayer based on membrane thickness and rigidity that best matches its TMD length (91-93). This model states that the molecular mechanism of TMD-dependent sorting among ER, Golgi, and plasma membrane is based on the fact that the membranes become thicker and more hydrophobic from ER toward the plasma membrane (91,93). This mechanism appears to play an important role in the retention of Golgi-resident proteins, such as the trans-Golgi resident enzyme β-1,4-galactosyltransferase(81), which contain a shorter membrane-spanning domain and does not exit the Golgi and becomes inserted into the cholesterol-rich plasma membrane. Similarly, another study has been performed on the ER-resident protein cytochrome B 5(84), which can translocate to the plasma membrane after increasing the length of its transmembrane domain. Although it is difficult to assign an exact length for ER retention, it might account for the targeting of Cosmc itself to the correct membrane.

The second model to explain ER localization of Cosmc focuses on the Cys residue within the TMD, which serves as a membrane localization signal. Palmitoylation is a post-translational modification in which the 16-carbon fatty acid is covalently attached to a protein, usually through a thio-ester linkage(94,95). The attachment of the 16-carbon fatty acid to integral membrane proteins occurs at a Cys residue located either within the transmembrane domain near to the cytoplasmic side or in the immediate vicinity of the cytoplasmic domain near the membrane. The demonstration of palmitoylation required for lysosomal enzyme-sorting of mannose-6-phosphate receptor has reinforced the hypothesis of the role of palmitoylation for normal trafficking and localization(96). However, while palmitoylation of the Cys residue in the TMD of this and other recycling receptors has been established to influence their correct membrane localization, there is no evidence that palmityoylation of Cys residues in the TMD of resident ER proteins can occur or that it is relevant to ER localization or retention. The palmitoyltransferases are not found in ER membranes and are commonly located on the plasma membrane and sorting vesicles derived from that membrane(97). Nevertheless, it is important in future studies to explore whether the single Cys in the TMD of Cosmc is modified with palmitate or some other lipid or substituent.

Our mutagenesis experiments with Cosmc showed that Cys19 mutation to Ser19 or Ala19 within the TMD caused the dramatically relocation of the protein from the ER to the Golgi. These results indicate that this cysteine residue within the TMD of Cosmc is very important for its ER localization. Previous studies have been done using a Golgi resident enzyme, galactosyltransferase (GT; UDP-galactose: beta-D-N-acetylglucosaminide β-1,4-galactosyltransferase)(95). A double mutant protein,

Cys29/His32 to Ser29/Leu32 within the TMD was found to be localized in the Golgi rather than the cell surface(95). It is also possible that the intermolecular disulfide bonds might form within the TMD at Cys19, and the di-sulfide-bonded protein might associate into dimers and non-covalent oligomers of the dimeric protein which may be too large to be incorporated into transport vesicles, and thus account for accumulation within the ER.

At present, our results show the important role of the TMD of Cosmc for its ER retention, but additional studies will need to be done to identify the role of Cys residue and other aspects of the mechanism for ER retention of Cosmc. The studies so far suggest that the sorting mechanism of resident ER membrane proteins could be more complex than a single model or retention motif. This work elucidates the critical region for localization of Cosmc and will be useful for further study of the chaperone role of Cosmc.

Chapter 3

Summary and Future Directions

Cosmc is a specific molecular chaperone for the activity of T-synthase, which catalyzes the reaction of addition of galactose onto the Tn antigen to form T antigen. Cosmc is a type II membrane protein, which has a short cytoplasmic domain, one transmembrane domain (TMD), and a long luminal domain. Previous studies have shown that Cosmc is mainly localized in the ER and functions as a molecular chaperone there. Therefore, which domain or region of Cosmc is required for its ER localization has become an important issue for us to explore. We used a traditionally well-known type II membrane protein, transferrin receptor which is localized to the plasma membrane in all the animal cells, and swapped each of the three domains with the corresponding domains in Cosmc to explore whether replacements lead to the localization of the new constructs to the ER. In this study, we mainly performed two experiments approaches to test our hypothesis. One is immunofluorescence staining, and the other is subcellular fractionation, which is used to confirm the results from the immunofluorescence staining. The data in this study have shown that the TMD of Cosmc is both necessary and sufficient for its ER localization. At the same time, we concluded that the cytoplasmic and luminal domains do not contain sufficient signals for ER localization, but they may have some accessory roles to help Cosmc target to the ER. Unexpected, we also found that the mutation of the lone cysteine residue within the TMD of Cosmc caused the loss of ER localization, which suggests that this single Cys is very important for the ER targeting of Cosmc.

The major future problem to address is to explore the possible mechanism as to how the TMD is involved in the ER localization of Cosmc. From previous studies, there are several models to be considered. First, considering the possibility that the length of the TMD could determine the final destination of a membrane protein, this seems unlikely, since a mutation of the Cys residue in the TMD caused mis-localization to the Golgi without altering the length. Nevertheless, it might be useful to explore the role of the size and composition of the TMD in the context of keeping or removing this important Cys residue that is conserved in all species expressing Cosmc homologs. It might be important to generate several new constructs, in which the TMD are extended by adding several extra leucines on both sides of the TMD to test whether increasing the length of TMD will lead to the loss of ER localization of Cosmc. Alternatively, the TMD could be shortened on each side of the key Cys residue. Second, it is possible that specific amino acids, such as Cys, within the TMD could be the keys to determine where Cosmc is sorted. To address this possibility, constructs should be made in which all residues flanking the Cys, and including the Cys, are mutated to appropriate alternative residues. Third, it is possible that the Cys residue is modified by post-translational modifications in some manner, such as lipid addition or co-valent interactions of disulfide bonds to other proteins that cause it retention in the ER. For this approach, Cosmc should be isolated from cells (or even organs such as liver) and studied by proteomic methods to identify the TMD sequence and identify possible modifications of the Cys or additional proteins co-purifying with Cosmc that may be disulfide bonded.

Our studies have established a new direction for studying ER localization mechanisms and imply the importance of specific residues in the TMD of a resident ER

protein. Future studies on Cosmc could lead to the identification of other novel ER proteins and helped to solve the mystery of ER localization versus Golgi targeting and targeting to more distal membranes in the secretory pathway. Successful studies in this direction may also be useful to aid in understanding the mechanisms of some related ER-associated protein-folding diseases and may provide new pharmacological targets for drug intervention.

References

1. Porter K.R., C. A., Fullam E.F. (1945) *J Exp Med.* **81**, 233-246
2. Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., and Darnell, J. (1999). In: Tenney, S. (ed). *Molecular Cell Biology*, 4th Ed., W. H. Freeman and Company, New York
3. Trowbridge, I. S., Collawn, J. F., and Hopkins, C. R. (1993) *Annu Rev Cell Biol* **9**, 129-161
4. Freedman, R. B. (1995) *Curr Opin Struct Biol* **5**(1), 85-91
5. Sidrauski, C., Chapman, R., and Walter, P. (1998) *Trends Cell Biol* **8**(6), 245-249
6. Roth, J. (1996) *Histochem Cell Biol* **106**(1), 79-92
7. Anfinsen, C. B. (1973) *Science* **181**(96), 223-230
8. Ellgaard, L., and Helenius, A. (2001) *Curr Opin Cell Biol* **13**(4), 431-437
9. Ellgaard, L., and Helenius, A. (2003) *Nat Rev Mol Cell Biol* **4**(3), 181-191
10. Frydman, J. (2001) *Annu Rev Biochem* **70**, 603-647
11. Sipe, J. D. (1992) *Annu Rev Biochem* **61**, 947-975
12. Muchowski, P. J. (2002) *Neuron* **35**(1), 9-12
13. Trombetta, E. S., and Parodi, A. J. (2003) *Annu Rev Cell Dev Biol* **19**, 649-676
14. Mayer, M. P., and Bukau, B. (2005) *Cell Mol Life Sci* **62**(6), 670-684
15. Herrmann, J. M., Malkus, P., and Schekman, R. (1999) *Trends Cell Biol* **9**(1), 5-7
16. Nakai, A., Satoh, M., Hirayoshi, K., and Nagata, K. (1992) *J Cell Biol* **117**(4), 903-914
17. Tasab, M., Batten, M. R., and Bulleid, N. J. (2000) *Embo J* **19**(10), 2204-2211

18. Tasab, M., Jenkinson, L., and Bulleid, N. J. (2002) *J Biol Chem* **277**(38), 35007-35012

19. DeBose-Boyd, R. A., Brown, M. S., Li, W. P., Nohturfft, A., Goldstein, J. L., and Espenshade, P. J. (1999) *Cell* **99**(7), 703-712

20. Willnow, T. E., Rohlmann, A., Horton, J., Otani, H., Braun, J. R., Hammer, R. E., and Herz, J. (1996) *Embo J* **15**(11), 2632-2639

21. Willnow, T. E. (1998) *Biol Chem* **379**(8-9), 1025-1031

22. Hsieh, J. C., Lee, L., Zhang, L., Wefer, S., Brown, K., DeRossi, C., Wines, M. E., Rosenquist, T., and Holdener, B. C. (2003) *Cell* **112**(3), 355-367

23. Shimizu, Y., and Hendershot, L. M. (2007) *Adv Exp Med Biol* **594**, 37-46

24. Branza-Nichita, N., Petrescu, A. J., Negroiu, G., Dwek, R. A., and Petrescu, S. M. (2000) *Chem Rev* **100**(12), 4697-4712

25. Hosokawa, S., Inui, M., Kajihara, Y., Matsuda, K., Ichitsubo, T., Pilgrim, W. C., Sinn, H., Gonzalez, L. E., Gonzalez, D. J., Tsutsui, S., and Baron, A. Q. (2009) *Phys Rev Lett* **102**(10), 105502

26. Wang, T., and Hebert, D. N. (2003) *Nat Struct Biol* **10**(5), 319-321

27. Hebert, D. N., and Molinari, M. (2007) *Physiol Rev* **87**(4), 1377-1408

28. Potter, B. A., Hughey, R. P., and Weisz, O. A. (2006) *Am J Physiol Cell Physiol* **290**(1), C1-C10

29. Spiro, R. G. (2002) *Glycobiology* **12**(4), 43R-56R

30. Fukuda, M. (2002) *Biochim Biophys Acta* **1573**(3), 394-405

31. Mall, M., Grubb, B. R., Harkema, J. R., O'Neal, W. K., and Boucher, R. C. (2004) *Nat Med* **10**(5), 487-493

32. Moncada, D. M., Kammanadiminti, S. J., and Chadee, K. (2003) *Trends Parasitol* **19**(7), 305-311

33. Kannagi, R. (2004) *Glycoconj J* **20**(5), 353-364

34. Brockhausen, I. (2006) *EMBO Rep* **7**(6), 599-604

35. Hennet, T. (2002) *Cell Mol Life Sci* **59**(7), 1081-1095

36. Colley, K. J. (1997) *Glycobiology* **7**(1), 1-13

37. Breton, C., Heissigerova, H., Jeanneau, C., Moravcova, J., and Imberty, A. (2002) *Biochem Soc Symp* (69), 23-32

38. Nilsson, T., Slusarewicz, P., Hoe, M. H., and Warren, G. (1993) *FEBS Lett* **330**(1), 1-4

39. Munro, S. (1995) *Embo J* **14**(19), 4695-4704

40. Qian, R., Chen, C., and Colley, K. J. (2001) *J Biol Chem* **276**(31), 28641-28649

41. Munro, S. (1995) *Biochem Soc Trans* **23**(3), 527-530

42. Munro, S. (1998) *Trends Cell Biol* **8**(1), 11-15

43. Opat, A. S., van Vliet, C., and Gleeson, P. A. (2001) *Biochimie* **83**(8), 763-773

44. Giraudo, C. G., and Maccioni, H. J. (2003) *Mol Biol Cell* **14**(9), 3753-3766

45. Bi, X., Corpina, R. A., and Goldberg, J. (2002) *Nature* **419**(6904), 271-277

46. Jaeken, J., and Matthijs, G. (2001) *Annu Rev Genomics Hum Genet* **2**, 129-151

47. Xia, L., Ju, T., Westmuckett, A., An, G., Ivanciu, L., McDaniel, J. M., Lupu, F., Cummings, R. D., and McEver, R. P. (2004) *J Cell Biol* **164**(3), 451-459

48. Xia, L., and McEver, R. P. (2006) *Methods Enzymol* **416**, 314-331

49. Ju, T., Cummings, R. D., and Canfield, W. M. (2002) *J Biol Chem* **277**(1), 169-177

50. Allen, A. C., Topham, P. S., Harper, S. J., and Feehally, J. (1997) *Nephrol Dial Transplant* **12**(4), 701-706

51. Ju, T., Brewer, K., D'Souza, A., Cummings, R. D., and Canfield, W. M. (2002) *J Biol Chem* **277**(1), 178-186

52. Yip, B., Chen, S. H., Mulder, H., Hoppener, J. W., and Schachter, H. (1997) *Biochem J* **321**(2), 465-474

53. Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M. A., Samyn-Petit, B., Julien, S., and Delannoy, P. (2001) *Biochimie* **83**(8), 727-737

54. Schneider, U., Schwenk, H. U., and Bornkamm, G. (1977) *Int J Cancer* **19**(5), 621-626

55. Ju, T., and Cummings, R. D. (2002) *Proc Natl Acad Sci U S A* **99**(26), 16613-16618

56. Ju, T., Aryal, R. P., Stowell, C. J., and Cummings, R. D. (2008) *J Cell Biol* **182**(3), 531-542

57. Nilsson, T., and Warren, G. (1994) *Curr Opin Cell Biol* **6**(4), 517-521

58. Rabouille C, N. T. (1995) *FEBS Lett.* **369**(No.1), 97-100

59. Barre, L., Magdalou, J., Netter, P., Fournel-Gigleux, S., and Ouzzine, M. (2005) *Febs J* **272**(4), 1063-1071

60. Zarei, M. M., Eghbali, M., Alioua, A., Song, M., Knaus, H. G., Stefani, E., and Toro, L. (2004) *Proc Natl Acad Sci U S A* **101**(27), 10072-10077

61. Munro, S., and Pelham, H. R. (1987) *Cell* **48**(5), 899-907

62. Scheel, A. A., and Pelham, H. R. (1996) *Biochemistry* **35**(31), 10203-10209

63. Townsley, F. M., Wilson, D. W., and Pelham, H. R. (1993) *Embo J* **12**(7), 2821-2829

64. Schutze, M. P., Peterson, P. A., and Jackson, M. R. (1994) *Embo J* **13**(7), 1696-1705

65. Jackson, M. R., Nilsson, T., and Peterson, P. A. (1993) *J Cell Biol* **121**(2), 317-333

66. Dahllof, B., Wallin, M., and Kvist, S. (1991) *J Biol Chem* **266**(3), 1804-1808

67. Fabre, E., Boelens, W. C., Wimmer, C., Mattaj, I. W., and Hurt, E. C. (1994) *Cell* **78**(2), 275-289

68. Cosson, P., and Letourneur, F. (1994) *Science* **263**(5153), 1629-1631

69. Harter, C., and Wieland, F. (1996) *Biochim Biophys Acta* **1286**(2), 75-93

70. Sato, K., Sato, M., and Nakano, A. (2003) *Mol Biol Cell* **14**(9), 3605-3616

71. Sato, M., Sato, K., and Nakano, A. (1996) *J Cell Biol* **134**(2), 279-293

72. Boehm, J., Letourneur, F., Ballensiefen, W., Ossipov, D., Demolliere, C., and Schmitt, H. D. (1997) *J Cell Sci* **110 (Pt 8)**, 991-1003

73. Hirata, T., Okabe, M., Kobayashi, A., Ueda, K., and Matsuo, M. (2009) *Biosci Biotechnol Biochem* **73**(3), 619-626

74. Otsu, M., Bertoli, G., Fagioli, C., Guerini-Rocco, E., Nerini-Molteni, S., Ruffato, E., and Sitia, R. (2006) *Antioxid Redox Signal* **8**(3-4), 274-282

75. Schweizer, A., Rohrer, J., Jeno, P., DeMaio, A., Buchman, T. G., and Hauri, H. P. (1993) *J Cell Sci* **104 (Pt 3)**, 685-694

76. Schindler, R., Itin, C., Zerial, M., Lottspeich, F., and Hauri, H. P. (1993) *Eur J Cell Biol* **61**(1), 1-9

77. Newton, T., Black, J. P., Butler, J., Lee, A. G., Chad, J., and East, J. M. (2003) *Biochem J* **371**(Pt 3), 775-782

78. Pfeffer, S. R., and Rothman, J. E. (1987) *Annu Rev Biochem* **56**, 829-852

79. Pelham, H. R. (1988) *Embo J* **7**(4), 913-918

80. Munro, S. (1991) *EMBO J.* **10**(No.12), 3577-3588

81. Nilsson, T., Lucocq, J. M., Mackay, D., and Warren, G. (1991) *Embo J* **10**(12), 3567-3575

82. Tang, Z. X., Chen, J. P., Sorensen, C. M., Klabunde, K. J., and Hadjipanayis, G. C. (1992) *Phys Rev Lett* **68**(20), 3114

83. Ronchi, P., Colombo, S., Francolini, M., and Borgese, N. (2008) *J Cell Biol* **181**(1), 105-118

84. Pedrazzini, E., Villa, A., Longhi, R., Bulbarelli, A., and Borgese, N. (2000) *J Cell Biol* **148**(5), 899-914

85. Rutledge, E. A., Gaston, I., Root, B. J., McGraw, T. E., and Enns, C. A. (1998) *J Biol Chem* **273**(20), 12169-12175

86. Rezaie, A. R., Fiore, M. M., Neuenschwander, P. F., Esmon, C. T., and Morrissey, J. H. (1992) *Protein Expr Purif* **3**(6), 453-460

87. Stearns, D. J., Kurosawa, S., Sims, P. J., Esmon, N. L., and Esmon, C. T. (1988) *J Biol Chem* **263**(2), 826-832

88. Fang, G., Weiser, B., Visosky, A., Moran, T., and Burger, H. (1999) *Nat Med* **5**(2), 239-242

89. Andersson, H., Kappeler, F., and Hauri, H. P. (1999) *J Biol Chem* **274**(21), 15080-15084

90. Sonnichsen, B., Fullekrug, J., Nguyen Van, P., Diekmann, W., Robinson, D. G., and Mieskes, G. (1994) *J Cell Sci* **107 (Pt 10)**, 2705-2717

91. Sprong, H., van der Sluijs, P., and van Meer, G. (2001) *Nat Rev Mol Cell Biol* **2**(7), 504-513

92. Bretscher, M. S., and Munro, S. (1993) *Science* **261**(5126), 1280-1281

93. Bulbarelli, A., Sprocati, T., Barberi, M., Pedrazzini, E., and Borgese, N. (2002) *J Cell Sci* **115**(Pt 8), 1689-1702

94. Dietrich, L. E., and Ungermann, C. (2004) *EMBO Rep* **5**(11), 1053-1057

95. Aoki, D., Lee, N., Yamaguchi, N., Dubois, C., and Fukuda, M. N. (1992) *Proc Natl Acad Sci U S A* **89**(10), 4319-4323

96. Schweizer, A., Kornfeld, S., and Rohrer, J. (1996) *J Cell Biol* **132**(4), 577-584

97. Husain, M., and Moss, B. (2003) *J Virol* **77**(16), 9008-9019