Diss. ETH No. 17445

Characterization of TMX4, a Novel ER Oxidoreductase with an unusual localization signal

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

DORIS ROTH

Dipl. Natw. ETH, ETH Zurich

born 27.01.1979

citizen of Germany

accepted on the recommendation of
examiner: Prof. Dr. Ari Helenius
co-examiner: Assoc. Prof. Dr. Lars Ellgaard
co-examiner: Prof. Dr. Hans-Peter Hauri
co-examiner: Prof. Dr. Yves Barral

2007

Seite Leer /
Blank leaf

Parts of this thesis will be published in the following article (manuscript included in the Appendix):

Roth, D., Lynes, E., Althaus, N., Simmen, T. and Ellgaard, L. (2007) A Non-classical Localization Signal Targets the Novel ER Oxidoreductase TMX4 to the Endoplasmic Reticulum. *Manuscript in preparation.*

Parts of the thesis presented here have been the result of collaborative efforts:

Nils Althaus, Institute of Biochemistry, ETH Zurich, performed the *in vitro* studies on the thioredoxin-like domain of TMX4 during his diploma thesis under my supervision. He also affinity-purified the peptide antibody against TMX4 and tested it in Western blots. Furthermore he investigated the membrane association of HA-TMX4 and conducted the initial topology studies with proteinase K.

Kathrin Roderer, Institute of Biochemistry, ETH Zurich, established biotinylation of cell surface proteins during her semester work under my supervision.

Seite Leer /
Blank leaf

TABLE OF CONTENTS

Summary

One single enzyme, protein disulfide isomerase (PDI), is able to catalyze both the introduction of disulfide bonds, their reduction and isomerisation in the endoplasmic reticulum (ER). However, almost 20 PDI-like thiol-disulfide oxidoreductases are known, indicating the complexity of the system that underlies the correct pairing of two cysteines into a native disulfide bond. To fully understand oxidative protein folding, it is necessary to elucidate the role of novel PDI-family members. In this thesis, we provide a first characterization of a novel ER thiol-disulfide oxidoreductase, TMX4 (thioredoxin-like transmembrane protein 4).

Sequence analysis showed that the protein contains a single thioredoxin-like domain, which holds an unusual CPSC active-site motif. The CPSC active-site sequence of HA-TMX4 is partially oxidized in the ER of living cells. Unlike other PDI-like family members it is predominantly oxidized, indicating that TMX4 has a different redox activity than well-known PDIs and that its *in vivo* redox state is subject to a different regulation. TMX4 is expressed in a wide variety of human tissues and is not upregulated by the unfolded protein response. By immunofluorescence, TMX4 was shown to localize to the ER. Furthermore, TMX4 was found to be a type I transmembrane protein that contains an Endoglycosidase H-sensitive *N*-glycan.

Interestingly, TMX4 does not have a classical ER retention signal at the C-terminus of its cytosolic tail. This region, however, holds an RQR sequence. Previously, RXR sequences have been found to confer ER localization to type II and multispanning membrane proteins. When the C-terminal tail was removed, HA-TMX4 partially redistributed to the cell surface. The same result was seen when mutating the two arginines of the RQR sequence to lysines, indicating that the RQR sequence contributes to the ER targeting of HA-TMX4. By attaching the RQR-bearing C-terminal region to a plasma membrane protein, CD4, we generated chimeras that remained in the ER. In the same construct, mutation of the RQR sequence resulted in cell surface expression. Overall, the results demonstrate that the RQR sequence in TMX4 is sufficient for ER localization.

Zusammenfassung

Ein Enzym, Proteindisulfidisomerase (PDI), ist ganz alleine in der Lage, Disulfidbrücken im endoplasmatischen Reticulum (ER) sowohl zu bilden, als auch zu reduzieren und isomerisieren. Dass dennoch beinahe 20 PDI-ähnliche Thioldisulfidoxidoreduktasen bekannt sind, weist auf die Komplexität hin, die der korrekten Paarung zweier Cysteine zu einer native Disulfidbindung zugrundeliegt. Um oxidative Proteinfaltung vollständig zu verstehen, ist es notwendig, die Rolle der neuen Mitglieder der PDI-Familie zu beleuchten. Im Rahmen dieser Dissertation legen wir die erste Charakterisierung einer neuen Thioldisulfidoxidoreduktase des ER, TMX4 (thioredoxin-like transmembrane protein 4), vor.

Die Analyse der Aminosäuresequenz von TMX4 zeigte, dass das Protein eine einzige Thioredoxin-ähnliche Domäne mit einem ungewöhnlichen CPSC-Motiv im aktiven Zentrum besitzt. Das ein CPSC-Motiv enthaltende aktive Zentrum von HA-TMX4 liegt in einem teilweise oxidierten Zustand vor. Anders als andere Mitglieder der PDI-Familie ist es vorwiegend oxidiert, was darauf hindeutet, dass sich die Redoxaktivität und die Regulierung des in *vivo* Redoxzustands von denen bislang charakterisierten PDI-ähnlichen Proteinen unterscheiden. Die Expression der TMX4-mRNA wurde in einer breiten Vielfalt menschlicher Gewebe nachgewiesen und nicht durch die "unfolded protein response", der zellulären Antwort auf ungefaltete Proteine, aufreguliert. TMX4 ist ein Typ-I-Transmembranprotein mit einem N-verknüpften Glykan, das sich gegenüber Endoglykosidase H sensitive verhält. Durch Immunofluoreszenz wurde gezeigt, dass TMX4 im ER lokalisiert ist.

Interessanterweise besitzt TMX4 kein klassisches ER-Lokalisierungssignal am C-Terminus des cytosolischen Schwanzes. Diese Region birgt jedoch eine RQR-Sequenz. Schon vor der vorliegenden Arbeit war bekannt, dass RXR-Sequenzen die ER-Lokalisierung von Typ-II-Transmembranproteinen und Proteinen mit mehreren Transmembrandomänen gewährleistet. Als der C-terminale Schwanz entfernt wurde, gelangte HA-TMX4 teilweise an die Zelloberfläche. Das gleiche Resultat wurde erhalten, als die Arginine der RQR-Sequenz zu Lysinen mutiert wurden, was darauf hindeutet, dass die RQR-Sequenz zum ER-"Targeting" von HA-TMX4 beiträgt. Durch Anheftung der RQR-beinhaltenden C-terminalen Region von TMX4 an das Plasmamembranprotein CD4 erzeugten wir Chimären, die im ER verblieben. Mutation der RQR-Sequenz im gleichen Konstrukt stellte die Zelloberflächen-

expression wieder her. Insgesamt zeigen die Resultate, dass die RQR-Sequenz von TMX4 hinreichend für die ER-Lokalisierung ist.

1 Introduction

1.1 Aim of this thesis

In the recent years, more and more protein disulfide isomerase (PDI) –like proteins became apparent. In a BLAST search for further PDI family members, genes of a number of previously uncharacterized proteins were identified. The work presented here focuses on one of these novel PDIs, called thioredoxin-like transmembrane protein 4 (TMX4). Based on its sequence, TMX4 belongs to the small subset of transmembrane proteins within the family and distinguishes itself from most family members by an unusual active site and the lack of a classical ER retention signal. The aim of this thesis was to characterize TMX4 on a molecular and cellular level with emphasis on redox properties and how ER localization is conferred.

1.2 Disulfide-bond formation takes place in the ER

The endoplasmic reticulum (ER) is the site of entry into the secretory pathway and a central organelle for protein folding and maturation. It is also the compartment where a number of co- and post-translational modifications like N-linked glycosylation, GPI-anchor attachment and disulfide-bond formation, i.e. the oxidation of two free cysteine thiols to form a covalent bond, take place (Ellgaard and Helenius, 2003). The ER lumen is more oxidizing than the cytosol and provides an appropriate and supportive environment for oxidative protein folding (Bass et al., 2004; Hwang et al., 1992). The correct formation of disulfides is critical during folding and for stabilization of the fully folded protein.

1.3 The PDI family of ER thiol-disulfide oxidoreductases

1.3.1 Thioredoxin-like domains and disulfide exchange

Although disulfide-bond formation does occur *in vitro* when oxygen and metals are present, it is a very inefficient process when not catalyzed by enzymes. In the ER, the oxidation, isomerization and reduction processes that ensure correct disulfide-bond formation are catalyzed by thiol-disulfide oxidoreductases of the protein disulfide isomerase (PDI) family (Ellgaard and Ruddock, 2005).

These enzymes contain at least one domain similar to thioredoxin (trx), a 12 kDa protein that functions as a disulfide reductase in the cytosol. The three-dimensional structure of thioredoxin and thioredoxin-like proteins shows a typical α/β fold with a central β-sheet surrounded by α-helices. The residues of the CXXC active-site motif (where X denotes any amino acid residue) are located at the N-terminus of the second α-helix and in the preceeding loop (Martin, 1995).

The PDI family members perform oxidation and reduction reactions by exchanging a disulfide with their substrates. The mechanism proceeds as followed: In case of an oxidase the two cysteines of the active site would be oxidized and transfer the disulfide bond to the substrate. Inversely, the cysteines of a reductase would be in a reduced state and attack the disulfide of the substrate. In detail, the N-terminal Cys (Cys_N) of the CXXC active site forms a thiolate group that attacks the sulfur atom of the substrate disulfide bond. It thereby breaks the original disulfide bond and forms an intermolecular disulfide bond with the attacked cysteine. This intermediate disulfide bridge is then released by the C-terminal cysteine of the CXXC motif. The reductase now carries the disulfide bond while the substrate is reduced. The intermediate disulfide linkage between enzyme and substrate is very short-lived and thus difficult to trap. To overcome this problem, CXXS active site mutants are better suited to identify reductase-substrate complexes because they lack the second cysteine that resolves the intermolecular disulfide bond.

1.3.2 Protein disulfide isomerase

The best characterized thiol-disulfide oxidoreductase in the ER is PDI itself (Goldberger et al., 1964). It is a remarkable enzyme that not only catalyzes oxidation, reduction and isomerization but also exhibits chaperone activity. PDI consists of four thioredoxin-like domains, referred to as **a**, **b**, **b'** and **a'**. Whereas **a** and **a'** each contain a CGHC active-site motif, the **b** and **b'** domains lack a functional active site. Nevertheless, **b'** plays an important role for the activity of PDI because it is indispensable for binding of non-native substrates and isomerase activity. Thus, in contrast to full-length PDI, the single **a** and **a'** domains are unable to catalyze isomerization, although they efficiently perform oxidation and reduction *in vitro* (Darby and Creighton, 1995; Klappa et al., 1995).

1.3.3 Novel members of the PDI family are emerging

Throughout the years, PDI has been the topic of most investigations in the field. In recent years though, many novel genes that encode PDI-like proteins have been identified. To date, a total of 17 such proteins have been found (appendix, table 1) (Ellgaard and Ruddock, 2005). The biological function of all these enzymes is currently an open question. Apart from possible differences in their redox activity, the enzymes may well act on different subsets of substrates that could be distinct or overlapping. For instance, the closest homolog of PDI, ERp57, has been found to act on only certain glycoproteins (Elliott et al., 1997; Oliver et al., 1997).

The function of the PDI-family members is not restricted to oxidative folding. For example, ERp44 regulates Ca^{2+} release by the inositol 1,4,5-trisphosphate receptor 1, a Ca^{2+} channel of the ER membrane, through a direct interaction that likely depends on the redox state of the receptor (Higo et al., 2005). It is also clear that some of the PDI-family members perform redox-unrelated functions, since they lack both cysteines of the active site. An example is ERp29 that binds thyroglobulin and enhances its secretion (Baryshev et al., 2006; Sargsyan et al., 2002). ERp29 is also important for transport across the ER membrane to the cytosol of murine polyoma virus during infection (Magnuson et al., 2005). Here, in a step potentially required for membrane penetration, ERp29 promotes partial unfolding of the major structural viral protein VP1.

The members of the PDI family differ in size, the number and arrangement of thioredoxin-like domains, and in part in tissue distribution (Ellgaard and Ruddock, 2005). Although most of the redox-active PDI-family members share a CXHC active-site sequence, there is considerable variation in the residues between the two cysteines. The nature of these two residues is known to influence the redox potential, i.e. the propensity of the active-site cysteines to be reduced or oxidized, and thereby the redox activity of a given enzyme (Grauschopf et al., 1995; Krause et al., 1991). Another interesting difference between the various enzymes is that some are soluble and others transmembrane-bound.

1.3.4 The TMX proteins of the PDI family

The transmembrane-bound subset of the family, designated as the 'TMX' proteins, includes the TMX protein, TMX2, and TMX3. All contain a single redox-active

thioredoxin-like domain but are otherwise not closely related. TMX has been shown to suppress apoptosis induced by brefeldin A when over-expressed in HEK293 cells (Matsuo et al., 2001). Whereas the knowledge about TMX2 is restricted to its cDNA sequence and tissue distribution (Meng et al., 2003), TMX3 is quite well studied *in vitro* in terms of structure-function relations (Haugstetter et al., 2005; Haugstetter et al., 2007). Nothing is known about substrate interactions of the TMX proteins but their membrane attachment could well influence these properties. As shown for the otherwise very similar ER chaperones, the soluble calreticulin and the transmembrane protein calnexin, the presence or absence of a transmembrane domain to a certain extent determines substrate binding (Danilczyk et al., 2000).

1.4 Arginine-based signals confer ER localization

In the crowded ER, a multitude of proteins with different destinations have to be discriminated from each other: whereas some proteins are to be secreted, unfolded or misfolded proteins must be retained. ER resident proteins should not escape the ER either and if they still do, they must be retrieved. Therefore, ER-localized proteins possess a signal that engages a specialized machinery for retention or retrieval. Such signals are per definition targeting signals if they function in a specific position in a protein, if they can be inactivated by mutation and if they confer ER localization to a reporter protein (Teasdale and Jackson, 1996). Two very well characterized examples of such signals in mammalian cells are the KDEL motif (Pelham, 1990) and the K(X)KXX sequence (Jackson et al., 1990; Nilsson et al., 1989).

1.4.1 The KDEL and K(X)KXX motifs are well-characterized ER retrieval signals

The extreme C-terminal tetrapeptide Lys-Asp-Glu-Leu (KDEL), is displayed by some soluble ER proteins such as chaperones and is both necessary and sufficient for their retrieval from post-Golgi compartments (Munro and Pelham, 1987; Pelham, 1988). Furthermore, when attached to the otherwise secreted lysozyme, the reporter protein remains localized to the ER (Pelham, 1989). The KDEL-dependent retrieval is executed by KDEL receptors, integral transmembrane proteins that reside in the ER Golgi intermediate compartment (ERGIC) and mainly in the Golgi apparatus (Griffiths et al., 1994; Lewis and Pelham, 1990; Lewis and Pelham, 1992b; Raykhel et al., 2007). Binding of KDEL receptors to their ligands, KDEL-tagged proteins, initiates retrograde

transport to the ER (Lewis and Pelham, 1992a), mediated by COPI-coated transport vesicles, a highly conserved transport machinery (Wieland and Harter, 1999).

COPI-coated carriers also carry out retrograde traffic of type I transmembrane proteins containing a discrete retrieval motif, K(X)KXX as shown *in vitro* (Cosson and Letourneur, 1994; Gaynor et al., 1994; Jackson et al., 1993; Letourneur et al., 1994; Orci et al., 1994; Presley et al., 1998). Like KDEL, the dilysine motif is a linear signal located at the C-terminus, which is in the cytosolic tail, and is responsible for retrieval (Jackson et al., 1990; Jackson et al., 1993; Nilsson et al., 1989). Unlike KDEL, however, the K(X)KXX sequence binds directly to a WD domains of either the α- or the β'-COP subunit (Eugster et al., 2004).

1.4.2 Regulation of K(X)KXX-mediated retrieval

A very interesting feature of the K(X)KXX motif was revealed by the discovery of its presence in some cell surface proteins as for example the high affinity IgE receptor α chain (Letourneur et al., 1995). Here, the K(X)KXX motif is used in quality control, retaining the α chain in the ER until the γ chain of the receptor correctly assembles with it. Upon binding, the K(X)KXX signal is masked and the complex can be targeted to the plasma membrane. Meanwhile, steric masking of the K(X)KXX motif to control complex formation has been recognized to be a widely used mechanism (Teasdale and Jackson, 1996).

Another regulatory aspect of the K(X)KXX motif is that the retrieval efficacy can be titrated by the sequence context (Itin et al., 1995). A weak dibasic retrieval motif may allow sorting proteins to post-ER compartments. Support for this idea was presented in studies about the protein ERGIC-53 which, with a KKFF motif, cycles between ER, ERGIC and the cis-Golgi and is not localized to the ER (Schweizer et al., 1988; Schweizer et al., 1990). Mutation of the motif could lead to accumulation in the ER (replacement of the phenylalanines by alanines) or to plasma membrane localization (substitution of the lysine at position -3 by an alanine) (Itin et al., 1995). Not only mutating the lysines disrupts their retrieval function, but also flanking them with glycines and prolines abolishes the ER targeting function. In contrast, a K(X)KXX motif can be active in a poly-glycine background as long as the lysines are directly surrounded by two serines (Teasdale and Jackson, 1996). These observations raised the question how the sequence context influences the retrieval efficacy. It is conceivable that it plays a role in exposure of the lysines or in whether the di-lysine

motif needs to attain a specific three-dimensional structure, as has been suggested for endocytosis motifs (Trowbridge et al., 1993), is unknown.

In addition to sequence context, the number of targeting signals per protein complex may influence the efficacy of retrieval, with more signals enhancing the efficacy. This was shown by comparison of chimeras between the K(X)KXX containing cytoplasmic tail of an adenoviral protein, E3/19K, and either CD4 or CD8. CD4 is a monomer, the chimera is almost exclusively co-localized with ERGIC-53, whereas the CD8 forms homodimers, and the chimera is mainly localized to the ER (Jackson et al., 1993).

1.4.3 Arginine-based signals

Both the KDEL and the K(X)KXX motif are well-characterized ER targeting signals. A less well characterized signal is a di-arginine motif which was first detected close to the N terminus of the invariant chain Iip35, a protein associated with the major histocompatibility complex class II (Schutze et al., 1994). As Ii is one of the rare type II membrane proteins of the ER, the Arg-based signal were first considered to be specific for type II membrane proteins. Meanwhile though, it has been found at various positions of polytopic membrane proteins and is recognized as a widespread motif.

The analysis of naturally occurring Arg-based signals and site-directed mutagenesis combined with a s screening approach (Zerangue et al., 2001) resulted in the proposal of a Φ/Ψ/R-R-X-R consus motif, in which Φ/Ψ and X represent an aromatic/bulky hydrophobic (often leucine) and any residue, respectively. More than two arginines and especially more copies of the whole motif render the localization particularly efficient. Although lysines and arginines are both basic amino acids and although the di-lysine and the Arg-based signals are similar in their sequence, lysines generally cannot substitute for arginines (Schutze et al., 1994; Zerangue et al., 1999). Yet, sequence context can determine if substitution of an arginine by a lysine is permissible (Teasdale and Jackson, 1996). Indeed, like K(X)KXX signal efficiency, the signal strength of the Arg signal can be influenced by the sequence context. Asparagines, negatively charged and non-polar residues, for example, weaken the motif when they occur immediately before the RXR or at the position of the middle X (Zerangue et al., 2001). As the K(X)KXX motif in ERGIC-53 mentioned above, Arg-based signals can be modulated to result in steady-state Golgi localization. In their

screening assay, Zerangue and colleagues (Zerangue et al., 2001) found signals of a broad range of strengths, including nonfunctional sequences. It is therefore recommendable to test the functionality of an Arg motif that is new or in an unusual context.

1.4.4 The spacing issue

The function of a di-lysine motif relies on strict spacing with respect to the C-terminus (Nilsson et al., 1989). The spacing requirements of Arg-based signals relative to the N-terminus in type II transmembrane proteins were therefore an important subject of investigation before it was recognized that Arg motifs were present in different cytosolic domains of multispanning membrane proteins (Michelsen et al., 2005). Moreover, signals were functional when appended to various positions in reporter proteins as long as they remained accessible (Zerangue et al., 2000). It turned out that for Arg-based signals, the distance to the membrane matters instead of the distance to the C terminus. Shikano and coworkers (Shikano and Li, 2003) even determined two different, only little overlapping zones where di-lysine and Arg motifs, respectively, are fully active: Arg signals function 16 – 46 Å away from the membrane whereas the K(X)KXX motif remains closer to the membrane.

1.4.5 Quality control checkpoints for oligomerization

Apart from the different distribution of di-lysine and Arg-based signals in type I and II transmembrane proteins, respectively, both motifs occur on a number of channels and receptors destined to leave the ER. In multimeric protein complexes, the functionality depends on correct assembly of the subunits. Unassembled subunits or incompletely assembled complexes expose Arg motifs and are retained in the ER, until the correct oligomerization - sometimes with different subunits combining to different heteromultimers (Michelsen et al., 2005) - leads to the masking of the Arg-based signal and allows forward transport to the plasma membrane (Zerangue et al., 1999). An example for subunits depending on Arg-based signals is the neuronal α4 acetylcholine receptor (O'Kelly et al., 2002). Well studied examples are also the inward rectifier potassium channel subunits (Kir) that assemble with four sulphonlylurea receptor proteins to octameric K_{ATP} channels in neurons or pancreatic

β-cells (Ma and Jan, 2002; Zerangue et al., 1999) and that have been used as model for studies on Arg-based signals. Kir6.2 carries an RKR sequence in its cytoplasmic tail that is responsible for ER localization and is masked by correct assembly of the K_{ATP} octamer. The coupling of subunit assembly to anterograde transport has also been investigated using the G-protein coupled γ-aminobutyric acid ($GABA_B$) neurotransmitter receptor where coiled-coil domains in the C-termini of both $GABA_B$ subunits mask Arg motifs not only in the context of the full-length protein but also when appended on a reporter protein (Margeta-Mitrovic et al., 2000).

1.4.6 The mechanism of ER localization via Arg-based motifs: retention or retrieval?

Since di-lysine and Arg-based motifs are both di-basic sequences, it seemed straightforward to propose that proteins with di-lysine and Arg-based motifs might be retrieved to the ER by the same cellular machinery. Indeed, with an avidity-enhancing tetrameric construct with the C-terminal tail of Kir6.2 subunits, it was possible to copurify COPI complexes from a 200,000 g pellet of precleared HeLa extract (Yuan et al., 2003). The COPI components specifically recognized the RKR sequence in the Kir6.2 tails. Moreover, the COPI subunit β-COP and protein segments containing Arg-based signals were found to interact in pull-down assays using immobilized peptides of either neuronal α4 acetylcholine receptor subunits or the invariant chain Iip35 terminus (O'Kelly et al., 2002).

The machinery that denotes ER localization signals is tightly linked to the choice of targeting mechanism: evidence for the involvement of the COPI coat would confirm a retrieval mechanism (instead of or in addition to a retention mechanism) and vice versa. The role of COPI complexes for Arg motif-mediated ER localization is therefore supported by the study of (Hardt et al., 2003) who demonstrated by glycan analysis that proteins with Arg signals are retrieved from early Golgi compartments. Still, in the question of retention versus retrieval via Arg-based signals, the case for retrieval is not made yet, as only a minority of Arg motif-bearing proteins show the glycan modification typical for early Golgi.

As for the function of COPI, many open questions remain since the interactions between COPI components and Arg-based signals have exclusively been

shown in vitro. Would COPI also recognize the Arg-based sequences in vivo? If so, which COPI subunit is the receptor for Arg-based signal? Is it not a COPI coat component at all and additional cellular machinery is involved? So far, only the binding of COPI to di-lysine signals was verified in a yeast two hybrid screen where notably all of the seven COPI component failed to bind a Arg-based motif (Zerangue et al., 2001). This result, together with the inconvertibility of di-lysine and Arg-based signals as well as their different spatial requirements (spacing to terminus and to the membrane, see above), suggests that in order to mediate ER localization, the two motifs engage a different mechanism, different machineries, different receptors that link them to the COPI-coated vesicles or at least different receptor binding pockets (Michelsen et al., 2005; Zerangue et al., 2001). There are also indications for a COPI-independent ER retrieval pathway (Girod et al., 1999).

1.4.7 Regulation of ER localization by steric masking

In order to act as checkpoints which ensure that only correct oligomeric assemblies exit the ER, Arg-based signal must be inactivatable once full oligomerization has been achieved. Steric masking is a main mechanism to inactivate the ER localization; steric masking by a partner subunit, described above, is the simplest release mechanism and implies that once a multimeric membrane protein is fully assembled, it is transported to the cell surface by default. In the light of recent discoveries, however, this might be a wrong concept (Ma and Jan, 2002). It seems that forward transport even of correctly assembled channels and receptors can be controlled by anterograde ER signals (Ma et al., 2001; Ma et al., 2002; Stockklausner et al., 2001). Furthermore, ER release of multimeric membrane proteins is affected by other regulatory binding proteins that mask ER localization signals. An additional layer of complexity arises from the finding that association between proteins with Arg-based signals and some of these regulator proteins is itself regulated by phosphorylation (Ma and Jan, 2002).

Scaffold proteins like Homers (Roche et al., 1999) and especially PDZ (PSD95/DLG/ZO-1) domain-containing proteins qualify for a function as such regulatory proteins. They have been shown to associate with channels and receptors, thereby promoting or inhibiting ER to Golgi transport (Scott et al., 2001; Standley et al., 2000; Xia et al., 2001). PDZ domain-proteins may allow cell surface expression in

the case of of the ionotropic N-methyl D-aspartate (NMDA) glutamate receptor. This receptor's subunit NR1-3 contains a PDZ domain-binding motif that suppresses ER localization conferred by an Arg-based motif in the NR1-1 subunit (Scott et al., 2001; Standley et al., 2000; Xia et al., 2001).

1.4.8 14-3-3 family members as regulatory proteins in forward transport

14-3-3 proteins constitute a family of proteins that are conserved in all eukaryotes investigated and stably assembled to homo- or heterodimers. Their exact function is unknown but they bind classical consensus motifs RSXS^{P}XP or RXXXS^{P}XP, where S^P denotes a phosphorylated serine (Tzivion and Avruch, 2002). They can act as adapters in protein-protein interactions and control the activity of enzymes. Signal transduction, cell-cycle control, apoptosis, stress response and malignant transformation are examples for the cellular processes, in which 14-3-3 proteins play a role (Muslin and Xing, 2000).

Interestingly, 14-3-3 proteins have also been demonstrated to influence subcellular localization of proteins. They function as regulatory proteins in post-Golgi sorting of receptors and channels and their clustering into membrane subdomains (Tzivion and Avruch, 2002; van Hemert et al., 2001). Therefore, 14-3-3 proteins became much discussed candidates for the function of suppressing Arg-based signals, when they were found to bind closely to Arg motifs and other di-basic ER localization signals in channels and receptors.

While screening *in vitro* for binding partners of a putative PDZ domain-binding sequence in KCNK3 potassium channels, (O'Kelly et al., 2002) identified two 14-3-3 proteins. The two isoforms of 14-3-3β interacted with a nonclassical, phosphoserine-containing sequence adjacent to the RK retention motif in the most C-terminal region of the KCNK3 subunits. In pull-down assays, 14-3-3β also recognized a series of other nonclassical motifs in nicotinic α4 acetylcholine receptor subunits and the Iip35 invariant chain (O'Kelly et al., 2002). The 14-3-3β protein and β-COP were shown to associate in a mutually exclusive way with KCNK3 subunits. The same could be the case for the acetylcholine receptor α4 subunit and the Iip35 invariant chain, because their Arg-based ER localization signal overlaps with the 14-

3-3 binding motif. COPI components also compete with 14-3-3 proteins for the tail of in the potassium channel subunit Kir6.2: two other 14-3-3 family members, 14-3-3ε and ζ, are recruited to the Arg-based motif in Kir6.2 in a screen with a tetrameric affinity construct of the C-terminus of Kir6.2 followed by an *in vitro* reconstitution experiment (Yuan et al., 2003).

Interestingly, the 14-3-3 proteins associate in vitro and in vivo with only dimeric or tetrameric Kir6.2 subunits, but not with monomeric ones. Since only the oligomeric Kir6.2 subunits were expressed on the plasma membrane, the binding of 14-3-3ε and ζ was correlated to the oligomeric state of the Kir6.2 tail (Yuan et al., 2003).

1.4.9 A model for how 14-3-3 proteins could probe the oligomerization of complexes

Taken together, these data suggest a novel role of 14-3-3 proteins in the ER-to-Golgi transport and a control mechanism has been proposed for correct oligomerization of multimeric complexes including membrane proteins: Not only correctly assembled multimers leave the ER, but also some unassembled subunits can escape. In the ERGIC and Golgi apparatus, the complexes and subunits are competed for by 14-3-3 proteins and COPI components because they bind in a mutually exclusive manner. The equilibrium depends on the valency and spatial arrangement of the signal-exhibiting tails or loops of the subunits. Correct oligomerization facilitates 14-3-3 binding and hampers association with the COPI coat, thereby promoting forward transport to the cell surface. When a complex is incorrectly or incompletely assembled, the subunits recruit COPI by their Arg signal and recycle to the ER in COPI vesicles (Nufer and Hauri, 2003; O'Kelly et al., 2002; Yuan et al., 2003).

As conceivable and catchy as the model is, there are several unresolved issues. Based on the data available, it cannot be ruled out that in addition to a retrieval-based mechanism, there is true retention where Arg motifs retain unassembled subunits in the ER by a yet obscure mechanism (Nufer and Hauri, 2003). Crucially, it is not known either where the Arg-based signals of the subunits actually are masked, in the ER, the ERGIC or the Golgi complex (Nufer and Hauri, 2003).

Moreover, it is still unclear how to implement the role of phosphorylation into the model. Given that their binding motifs contain phosphoserine, 14-3-3 proteins can

be viewed as cofactors of cellular regulation by serine/threonine phosphorylation (Tzivion and Avruch, 2002). Hence it was not a surprise to detect that serine phosphorylation controlled the interaction of 14-3-3β with KCNK3 channels and the forward transport of Iip35 with MHC class II molecules. Forward transport was thus suggested to be initiated by phosphorylation of 14-3-3 binding motifs and subsequent binding by 14-3-3 family members. Association with Kir6.2, however, was independent of phosphorylation. In addition, negatively charged amino acids have been applied to mimic phosphoserines (Scott et al., 2003; Scott et al., 2001; Standley et al., 2000; Xia et al., 2001). These substitutions, however, do not necessarily show that the observed effects happen as a result of an *in vivo* phosphorylation event.

The pivotal role of COPI in the model is a further problem because the role of COPI in Arg-based ER localization has not yet been convincingly shown (as described above), especially due to the absence of in vivo data. COPI binding is, however, not the only association event that needs in vivo verification. Also 14-3-3 binding must still be tested - in vivo and in the context of the entire multimeric membrane complexes, not only using the subunit tails used in the in vitro approaches (Nufer and Hauri, 2003).

1.4.10 Arg-based signals in other proteins than channels or receptors

Although Arg-based signals in channels, receptors and other protein complexes (like the Iip35 invariant chains associated with MHC class II molecules) have up to now been focused on most, Arg signals and their feature to act as (transient) ER localization signals may also be used for other functions than coupling assembly state to anterograde transport.

Coupling of the folding status to anterograde transport of a membrane protein has also been suggested. When Arg-based motifs were found in the cystic fibrosis transmembrane conductance regulator (CFTR) protein (Chang et al., 1999), it was thought that transient exposure of Arg motifs of improperly folded proteins results in ER localization. This idea was based on mutational studies of the arginines in the Arg-based signals. Substituting the arginines lead to cell surface exposure of the ΔF508 mutant form of the protein, which is normally degraded by ER-associated degradation and which is the cause of disease in most cystic fibrosis patients

(Michelsen et al., 2005). It seems, however, as though the loss of an ER exit signal is responsible for ER retention of the ΔF508 mutant, not the exhibition of a usually shielded Arg motif (Wang et al., 2004). The absence of one of the mutated arginine signals in mouse CFTR also weakens the concept of the role that Arg-based signals play for CFTR. Still, that Arg motifs may affect the fate of the ΔF508 mutant, raised interest in their influence in disease-causing mutations, like SUR1 mutations that correlate with hyperinsulinism (Cartier et al., 2001) and vasopressin receptor 2 mutations that correlate with diabetes insipidus (Hermosilla et al., 2004).

An unexpected presence of Arg motifs was observed in type I transmembrane proteins of viruses like Epstein-Barr virus, human cytomegalovirus, and herpes simplex virus type I (Lee, 1999; Meyer et al., 2002; Meyer and Radsak, 2000). In these cases, Arg-based signals mediate access to the inner nuclear membrane, which is usually not granted to resident proteins of the secretory pathway, but contains a distinct set of integral proteins (Rolls et al., 1999). During the maturation of infectious particles, viral transmembrane glycoproteins are translocated from the site of their synthesis into the inner nuclear membrane, before DNA-containing nucleocapsids are enveloped (Roizman and Furlong, 1974). Neither the mechanism of the translocation of the viral proteins to the inner nuclear membrane, nor the precise function of the viral type I membrane proteins are known (Meyer and Radsak 2000). It is interesting to note, though, that Arg-based signals are possibly involved in targeting certain proteins to a specialized membrane domain of the ER.

Although Arg-based signals on type I transmembrane proteins seem uncommon, they are not only found in viral "specimens" but also on the animal lectin VIPL, an ER-resident protein (Nufer et al., 2003).

2 Results

2.1 Bioinformatics analysis

2.1.1 Sequence analysis of TMX4

The human gene encoding TMX4 (EnsEMBL geneID: ENSG00000125827) was identified in a BLAST search using a consensus sequence for a thioredoxin-like domain to query GenBank at the NCBI website (www.ncbi.nlm.nih.gov). The open reading frame predicts a protein of 349 residues (39 kDa, pI=4.3) including a signal sequence for entry into the secretory pathway with the signal peptide peptidase cleavage site predicted between residues 23 and 24 (Fig. 1). The thioredoxin-like domain covers residues 39 to 146 and is followed by a stretch of 41 residues preceeding a predicted transmembrane helix spanning residues 188-210 (see also Fig. 4B). The C-terminal region (residues 211-349) is highly negatively charged with 54 acidic residues and contains no classical ER-retention motif of the K(X)KXX type often found in type I transmembrane proteins of the ER.

2.1.2 Phylogenetic analysis of the TMX4 and TMX sequences

Sequence alignments with members of the human PDI family revealed that TMX4 is closely related to TMX, in particular in the thioredoxin-like domain where the two proteins show 53% sequence identity. In contrast, the C-terminal region beyond the predicted transmembrane helix is considerably less well conserved, with the most obvious difference being the length (78 residues in TMX versus 139 residues in TMX4). Further sequence analysis showed that both proteins are present in a variety of mammals, amphibia, fish and birds but not in plants and yeast.

The active site sequence does not seem to be the most important determinant for defining the protein identity because some TMX4 homologs, notably those of more distant organisms to human, have a CPAC motif like TMX. This sequence is also used in those species – among them *C. elegans* - that interestingly only have one proteins that is orthologous to both proteins. We were particularly interested in whether the *C. elegans* protein DPY-11 corresponds to TMX4 because dpy-11 had already been characterized (Ko and Chow, 2002).

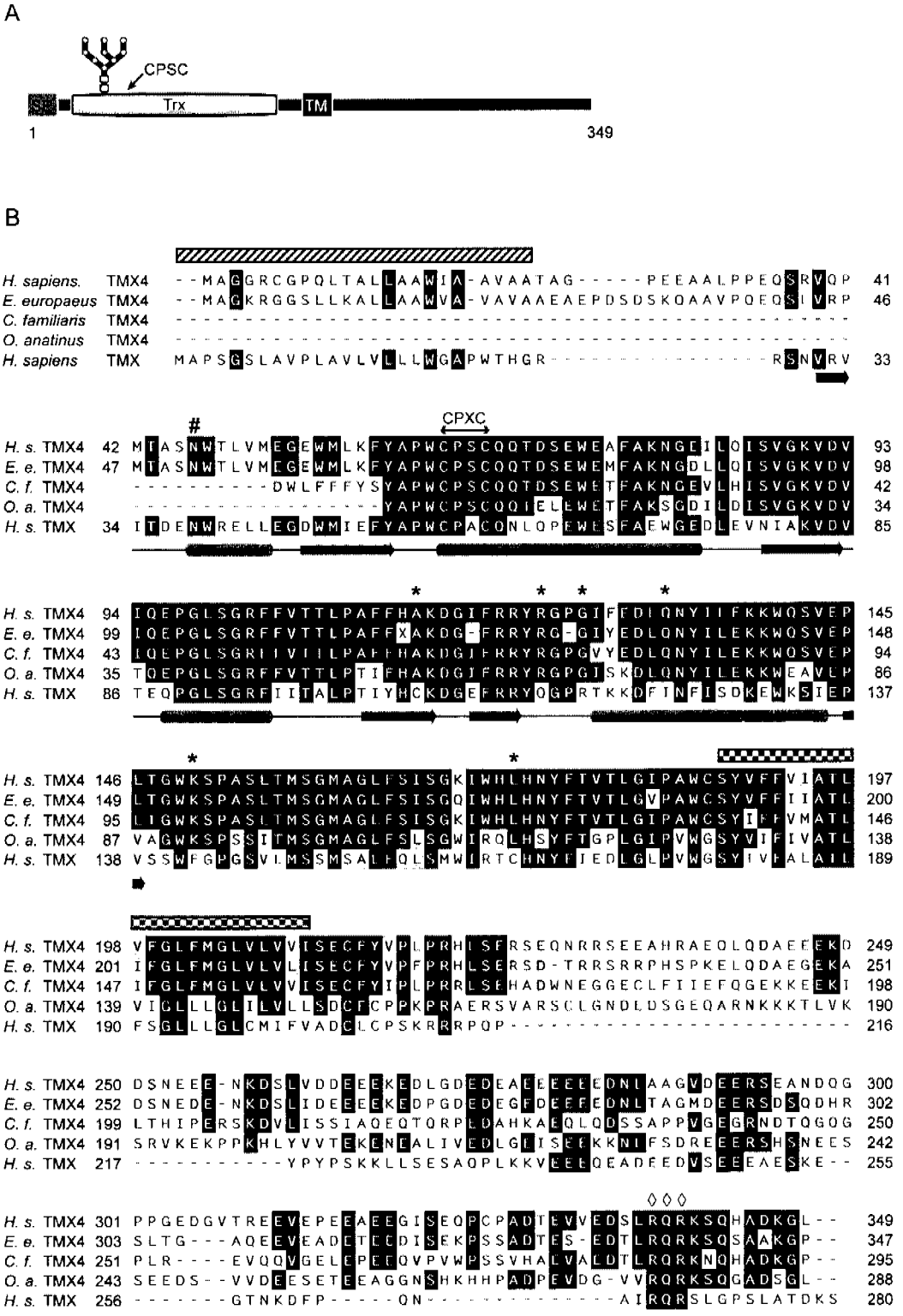


FIG. 1. **Sequence features and evolutionary conservation of TMX4.** A, graphical overview of TMX4 (Ensembl Gene ID ENSG00000125827). The following sequence features are indicated: predicted signal peptide (SP), N-glycosylation consensus site (schematic drawing of a glycan), thioredoxin-like domain (Trx), active-site motif (CPSC) and predicted transmembrane domain (TM). B, multiple sequence alignment of human TMX4, its orthologs in *Erinaceus europaeus* (Ensembl Gene ID ENSEEUG00000014678), *Canis familiaris*

(Ensembl Gene ID ENSCAFG00000005883) and *Ornithorhynchus anatinus* (Ensembl Gene ID ENSOANG00000008415) and its human paralog TMX (Ensembl Gene ID ENSG00000139921). Black boxes represent amino acid identities and gray boxes show amino acid similarities. Secondary structure elements of the TMX thioredoxin-like domain as determined by NMR spectroscopy are added below the TMX sequence, with cylinders and arrows representing α-helices and β-strands, respectively (Tochio et al.).

To learn more about the evolutionary relationship between these proteins, we subjected all TMX and TMX4 orthologs available on Ensembl to phylogenetic analysis. The sequences were aligned with ClustalW XXL and phylogenetic analysis was performed by PhyML (Guindon and Gascuel, 2003) using the maximum likelihood method, and plotted with PHYLIP (Felsenstein, 1989).

The resulting evolutionary tree is shown in Fig. 2. By including ERp57 (not shown in the figure), we rooted the tree so that the direction of the path indicates the passage of evolutionary time. The TMX/TMX4 proteins cluster in three main groups. One cluster comprises the organisms that have only one TMX/TMX4 version. This cluster, including *C. elegans* DPY-11, diverged first from a common TMX/TMX4 ancestor. After that, this ancestor duplicated to give rise to the paralogs TMX and TMX4 which evolved along separate paths and cluster in the evolutionary tree in two separate groups. Between TMX and TMX4, it is TMX that is closer related to the *C. elegans* protein dpy-11, because the branches in the TMX cluster are shorter than the TMX4 branches. As branch length represents the number of sequence changes that occurred prior to the next level of separation, the longer branches in the TMX4 cluster indicate that TMX4 acquired more mutations and diverged more from the common TMX/TMX4 ancestor. Human TMX, however, could not rescue a dpy-11 mutation (Ko and Chow, 2002). Taken together, the *C. elegans* dpy-11 protein corresponds neither to TMX nor to TMX4.

2.1.3 Sequence comparison between TMX4 and TMX

With the aid of the phylogenetic analysis and multiple sequence alignments we could now more easily analyze the differences between the two closely related proteins and pinpoint residues in the N-terminal 187 residues that were defining for the TMX4 and TMX proteins, respectively. We identified six such residues that were (with very few exceptions) conserved between species for each paralog but different in the other

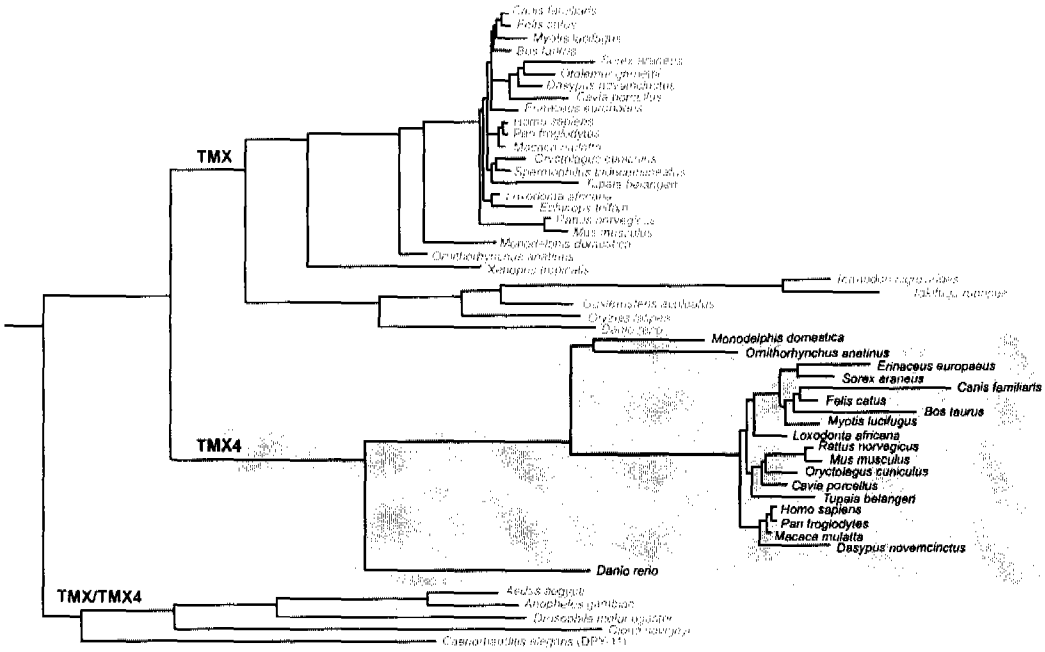


FIG. 2. **Phylogenetic tree for TMX and TMX4.** Amino acid sequences of TMX and TMX4 of all organisms available from all organisms available were obtained from the Ensembl genome browser (release 45, June 2007). Full-length sequences were aligned on the ch.EMBNET.org server, using ClustalW XXL. The alignments were carried out using the blosum series matrix with an open gap penalty of 10 and an extended gap penalty of 0.05. Phylogenetic trees were constructed using PhyML and Phylip. Human ERp57 was selected as outgroup. The TMX4 cluster is highlighted in grey. Separate branches are labeled with the name of the individual organisms. Purple, TMX; black, TMX4; green, organisms with only one version of the protein.

(TMX4 residue/TMX residue): Ala114/Cys106, Arg123/Gln115, Gly126/Arg118, Gln132/Ile124, Lys150/Phe142 and Leu173/Cys165. Variances of these mostly conserved residues are the positions of Ile 124 and Phe142 of TMX which are occupied by a different large hydrophobic amino acid in some organisms. Furthermore, the Leu173/Cys165 pair show slight variations with an isoleucine instead of Leu173 in TMX4 and a serine at the position Cys165 of TMX. The Lys150/Phe142 and Leu173/Cys165 residues are found in the region between the thioredoxin-like domain and the predicted transmembrane helix. Since no structural information exists for this region it is difficult to assess possible implications of these substitutions although it is worth noticing the exchange of a Leu in TMX4 for a Cys in TMX. The first four residues are present in the thioredoxin-like domain. To analyze

potential functional consequences of amino acid substitutions at these positions we turned to the NMR structure of the TMX thioredoxin-like domain (PDB accession number: 1X5E; Fig. 3A) that has been solved as part of the RIKEN Structural Genomics Initiative (Tochio et al., 2005).

This NMR structure was solved for a fragment corresponding to residues 39-150 of human TMX4 and showed a canonical thioredoxin fold with a β_1-α_1-β_2-α_2-β_3-α_3-β_4-β_5-α_4 arrangement of secondary structure elements followed by an additional α-helix and a short β-strand (Fig. 1B; Fig. 3A). For comparison, we built a model of TMX4 with residues 25-144 as input using the Robetta Structure Prediction Server (www.robetta.org) (Fig. 3B) (Chivian et al., 2005).

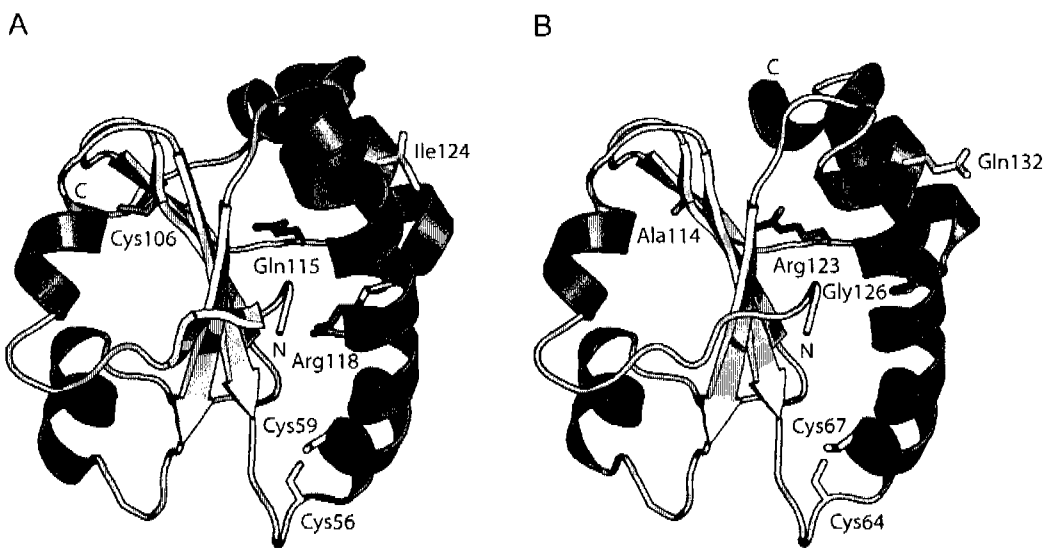


FIG. 3. **Structural features of TMX and TMX4.** A, NMR structure (conformer #1) of TMX determined by the RIKEN Structural Genomics Initiative (Tochio et al., 2005). The fragment comprising residues 30-139 is depicted. B, model of the TMX4 thioredoxin-like domain generated using the Robetta server (www.robetta.org) with residues 25-144 as input (Chivian et al., 2005). In the two representations α-helices are shown in red, β-strands in yellow and loops in green. Sidechains of individual residues are depicted as stick models and color coded so that corresponding residues in the two proteins have the same colors. Apart from the active-site cysteines (cyan), the four residues shown are those that define the two proteins (see main text for details): Cys106/Ala114 (purple), Gln115/Arg123 (hot pink), Arg118/Gly126 (dark blue), and Ile124/Gln132 (pink) (TMX residue/TMX4 residue). The N- and C-termini are labeled. The representations were generated using Pymol (DeLano, W.L. The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA, http://pymol.sourceforge.net/).

We were specifically interested in analyzing the four residues in the thioredoxin-like domain mentioned above that were not conserved between TMX4 and TMX throughout species. Cys106 (Ala114 in TMX4) is located in β_4 with its side chain completely buried and packing against that of Phe111, which is conserved in TMX4 (Phe119). So, we expect no structural or functional effects of this substitution. Ile124 is located in the middle of α_4, and its side chain mostly contacts that of Trp72 in α_2. In TMX4 these two hydrophobic residues are substituted by the polar Gln132 and Asn80. Hydrogen bonding between these residues might well provide the stability that is offered by the hydrophobic interaction of Ile124 and Phe111 in TMX. Arg118 in TMX is an interesting residue because most thioredoxin-like domains of human PDI family members contain an Arg at the corresponding position. This residue, positioned in the loop between β_5 and α_4, has been proposed to move in and out of the active site, thereby modulating the pK_a of the C-terminal cysteine in the CXXC motif by means of its charge (Lappi et al., 2004). Curiously, this Arg in TMX is substituted by Gly126 in TMX4. Instead, we found a conserved Arg in position 123 of TMX4 whereas TMX has a conserved Gln at this position. It is tempting to speculate that Arg123 performs the same function in TMX4 as Arg118 does in TMX.

2.2 *TMX4 is a protein of the ER*

Next, we wanted to verify experimentally the predicted ER localization of TMX4. To do so, we first expressed TMX4 with an HA epitope inserted right after the predicted signal peptidase cleavage site to generate the N-terminally HA-tagged protein, HA-TMX4, in Vero cells. This monkey cell line is particularly well suited for immunofluorescence microscopy because the cells are flat and large enabling a clear visualization of the ER. The cells were fixed according to the Nakane protocol (McLean and Nakane, 1974) and permeabilized with saponin before they were stained with primary antibodies against the ER protein calnexin (CNX) and the HA tag, and then fluorescently labeled secondary antibodies. The results showed that HA-TMX4 largely colocalized with CNX (Fig. 4A).

Although HA-TMX4 showed the typical reticular structure of the ER, we tested whether it also localized to the Golgi apparatus and the ERGIC in Vero cells. HA-TMX4 did not colocalize with the Golgi marker giantin (data not shown). To

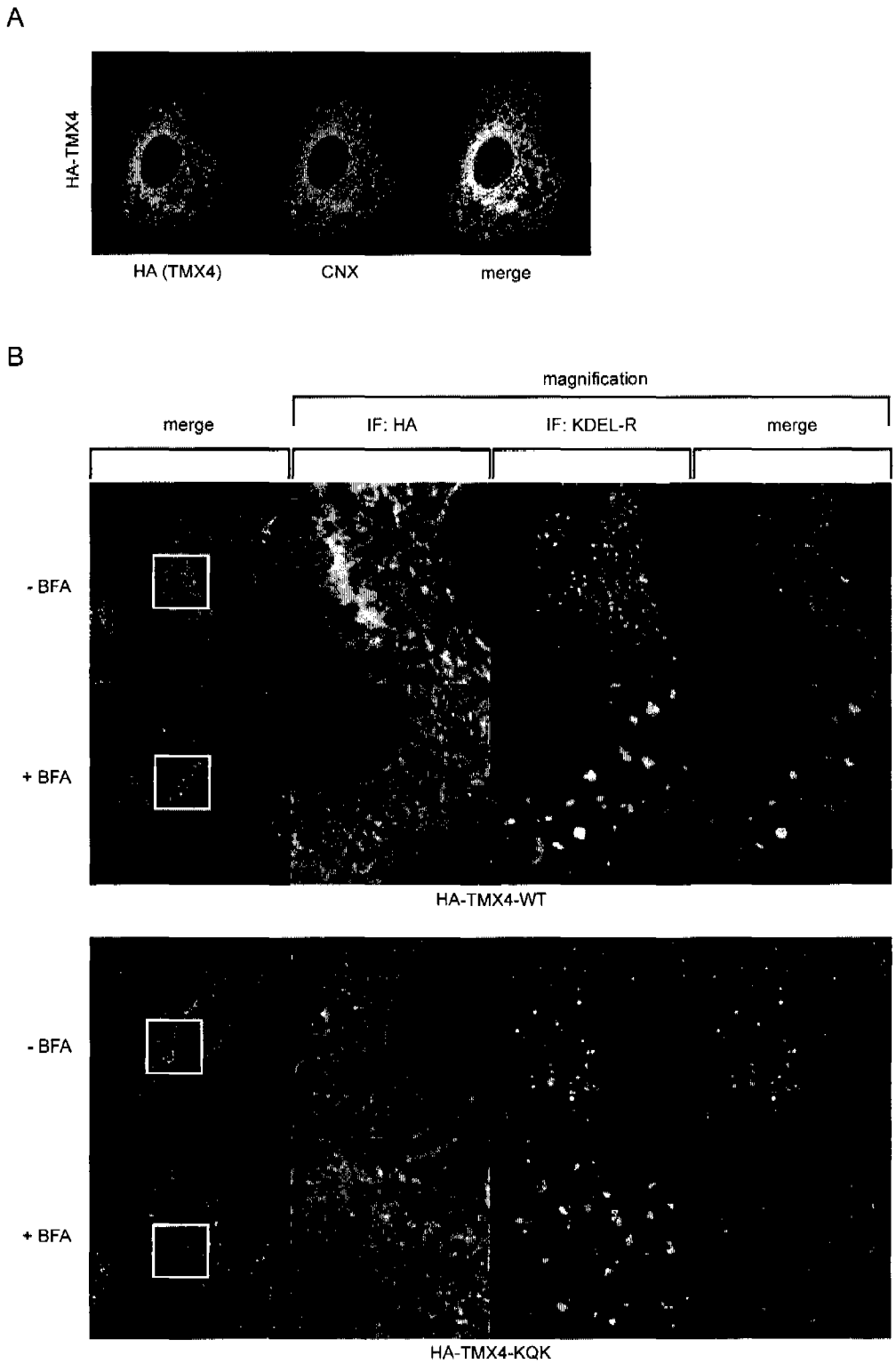


FIG. 4. **TMX4 is an ER-resident protein.** A, HA-TMX4 colocalizes with the ER-marker calnexin (CNX). Vero cells were transfected with HA-TMX4, fixed and analyzed by immunofluorescence using antibodies against the HA epitope and CNX. B, HA-TMX4 and HA-TMX4-KQK do not localize with marker protein that rapidly cycles in the ERGIC. Vero cells were transfected with HA-TMX4 (upper panel) and HA-TMX4-KQK (lower panel) prior to mock- or BFA treatment. Cells were then fixed and processed for

immunofluorescence using antibodies against HA analyzed by immunofluorescence using antibodies against the HA epitope and against the KDEL receptor (KDEL-R). The white rectangles in the first picture of each row indicate the area that was magnified in the following pictures.

better visualize a potential ERGIC colocalization, we mock-treated the cells or incubated them with brefeldin A (BFA) (Fig. 4B). BFA can be used to determine if proteins rapidly cycle between ER and post-ER compartments. Proteins that do so, such as the KDEL receptor (Tang et al., 1995) and ERGIC-53 (Lippincott-Schwartz et al., 1990) accumulate in dotted structures upon treatment with the drug. In such dotted structures, HA-TMX4 WT and KQK did not colocalize with the KDEL receptor (Fig. 4B). This result indicates that HA-TMX4 is not a rapidly cycling protein in the ERGIC.

2.3 Membrane topology and glycosylation status of TMX4

The TMX4 sequence contains one consensus site for *N*-glycosylation on Asn46 (Fig. 1B). To investigate whether HA-TMX4 is *N*-glycosylated, we expressed the protein is HeLa cells and treated the lysate with endoglycosidase H (Endo H), an enzyme that removes high-mannose sugars typically found on ER proteins. The subsequent Western blot analysis showed that HA-TMX4 indeed contained one glycan sensitive to Endo H treatment (Fig. 5A).

The TMX4 sequence was analyzed for the presence of transmembrane regions using the TMHMM algorithm (Sonnhammer et al., 1998), which predicted TMX4 to be a type I transmembrane protein with residues 188-210 spanning the membrane (Fig. 5B). In Fig. 5C we established that TMX4 is indeed an integral membrane protein. Isolated crude membranes from HA-TMX4-expressing HeLa cells were extracted with sodium carbonate at high pH, and soluble and membrane-associated proteins were separated by high-speed centrifugation. As controls for a membrane and a soluble protein we analyzed the distribution of CNX and ERp57, respectively. When probing the two fractions by Western blotting the soluble ER protein ERp57 was partly detected in the membrane fraction. Previously, we have obtained a similar result for another soluble ER protein, calreticulin (Haugstetter et al., 2005), a finding that is most likely due to the association of both ERp57 and calreticulin with membrane proteins. In contrast, HA-TMX4 was detected solely in the insoluble

fraction, as was the case for the control protein CNX. In addition to the transmembrane domain two other hydrophobic stretches were detected in TMX4 by the TMHMM algorithm: The first, at the very N-terminus, corresponded to the signal peptide and the second comprised residues 150-187, i.e. the region that links the thioredoxin-like domain to the transmembrane domain. A priori, we could not rule out that this linker region was also be buried in the membrane, thereby creating a topology with both termini of TMX4 in the ER and a short cytosolic loop. We therefore went on to determine the membrane topology of TMX4 using a proteinase K protection assay (Fig. 5D).

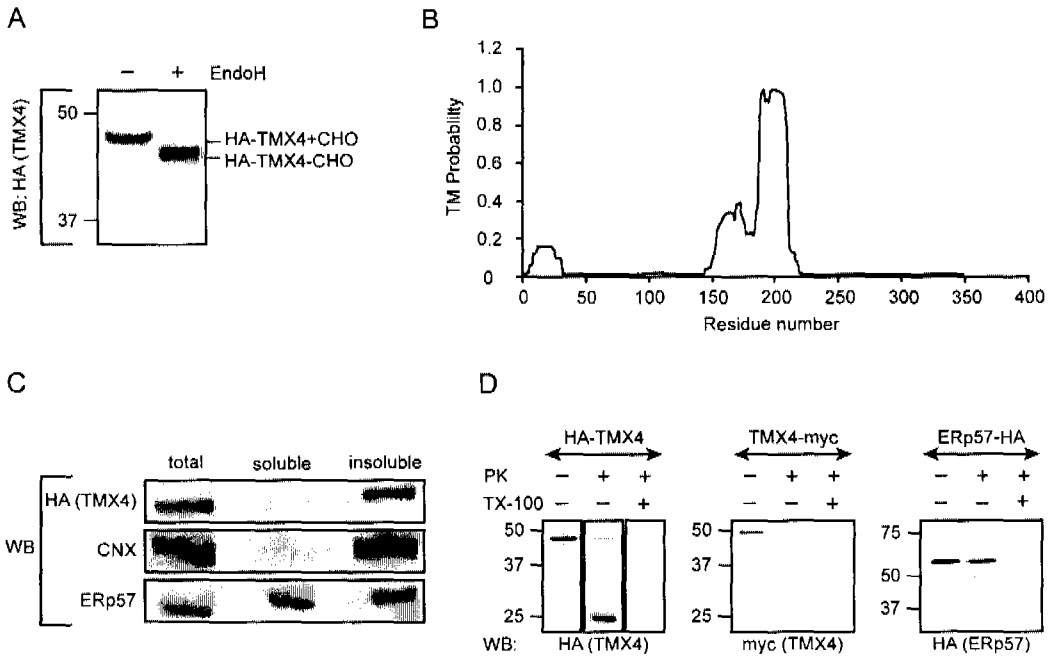


FIG. 5. **TMX4 is an N-glycosylated type I transmembrane protein.** A, HA-TMX4 is endoglycosidase H- (Endo H-) sensitive. Lysates from HeLa cells transiently expressing HA-TMX4 were mock treated or treated with Endo H and subjected to SDS-PAGE before Western blotting with an HA-specific antibody. B, Transmembrane (TM) topology prediction of the TMX4 sequence. The TMX4 amino acid sequence was analyzed by the transmembrane hidden Markov model (TMHMM) and the probability for the individual residues to be in a transmembrane helix were plotted versus the residue number. The stretch between residues 5 and 35 constitutes the predicted signal sequence. C, HA-TMX4 is an integral membrane protein. After subjecting crude membranes of HA-TMX4-transfected HeLa cells to alkali extraction, the soluble and insoluble fractions were separated by ultracentrifugation through a sucrose cushion. The distributions of HA-TMX4 and the marker proteins CNX (a type I transmembrane protein) and ERp57 (a soluble ER protein) were visualized by Western blotting. D, TMX4 is a type I transmembrane protein. HeLa cells were transfected as indicated and metabolically labeled. Crude membranes were isolated and mock treated or treated with Proteinase K in the presence or absence of 1% Triton X-100 (TX-100). The digestion products were isolated by immunoprecipitation with an anti-HA and anti-myc antibody, respectively, before they analyzed by SDS-PAGE and visualized by autoradiography.

First, HeLa cells were transfected with N-terminally HA-tagged TMX4 and metabolically labeled overnight. Crude membrane extracts were then left untreated or treated with proteinase K thereby digesting polypeptide chains not protected by the ER membrane. After immunoisolation, the eluates were analyzed by SDS-PAGE and autoradiography. The untreated sample gave a signal at the expected size for HA-TMX4 of ~45 kDa. Upon proteinase K treatment the band shifted to 25 kDa, a size that closely matches that of the 21 kDa expected for ER luminal and transmembrane domain of HA-TMX4. When the ER membrane was solubilized by addition of the detergent Triton X-100, proteinase K gained access to the ER lumen and the signal disappeared. These results showed that the HA-tagged N-terminus is localized in the ER as expected based on the presence of the N-glycan on Asn46. They did not show, however, whether the C-terminus was in the cytosol or in the ER. Thus, we repeated the proteinase K assay with HeLa cells expressing C-terminally myc-tagged TMX4. When the sample was treated with proteinase K, the TMX4-myc signal disappeared almost completely and at the same time no smaller immunoreactive band appeared, strongly indicating that the myc-tagged C-terminus was not protected inside the ER. As a control for ER membrane integrity, we performed the proteinase K experiment on membranes from HeLa cells transfected with HA-ERp57. Since the ERp57 signal was not affected by proteinase K treatment unless detergent was added, we conclude that the ER membranes were intact during the experiment.

Taken together, the results demonstrated that TMX4 is an *N*-glycosylated type I transmembrane protein, localized in the ER.

2.4 TMX4 does not contain a structural disulfide bond

Apart from the two active-site cysteines, TMX4 contains cysteines at positions 187, the very last residue before the transmembrane region, as well as 213 and 326 in the cytosolic tail. Although unlikely, we could not completely rule out that the two cytosolically localized cysteines form a disulfide bond. Moreover, the single cysteine on the lumenal side might be engaged in an intramolecular disulfide bridge. We therefore incubated HA-TMX4-expressing HeLa cells with N-ethylmaleimide (NEM) to preserve disulfide bonds, before we lysed the cells in loading buffer (see Experimental procedures) without DTT and resolved the lysates on SDS-PAGE under nonreducing conditions. HA-TMX4 showed the same migration as reduced control

samples (data not shown). We therefore concluded that HA-TMX4 does not form structural disulfides or disulfide-linked dimers with protein partners, while we cannot exclude covalent bonds to low molecular weight compounds such as reduced glutathione.

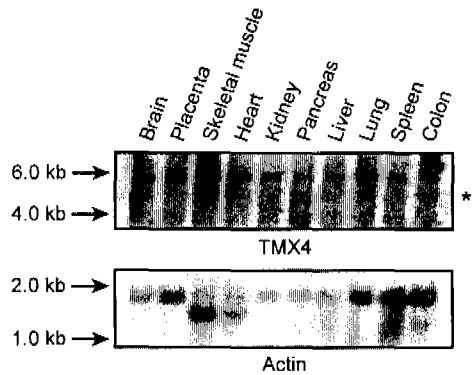
2.5 Transcriptional analysis of TMX4

To investigate the tissue-distribution of TMX4, we investigated the transcript level in a variety of human tissues by Northern blotting (Fig. 6A). To this end, ^{32}P-labeled TMX4 probes were used to hybridize a commercial tissue Northern blot. By autoradiography, we detected a signal at the predicted size of 6.0 kb for the TMX4 transcript (Ensembl) in all tissues tested, although with varying intensity. The highest transcript levels were found in brain, kidney, heart, skeletal muscle and lung. The signal for TMX4 was somewhat weaker than the 2.0 kb actin signal used as a control, that showed the expected presence of an additional isoform at 1.8 kb in heart and skeletal muscle.

TMX4 is a thiol-disulfide oxidoreductase and therefore expected to have a function in protein folding, specific quality control or coping with ER stress. Therefore, we tested whether TMX4 is upregulated in the unfolded protein response (UPR), one of the cellular answers to alleviate ER stress (Fig. 6B). We transfected HeLa cells with HA-TMX4 before exposing the cells to the two UPR inducers tunicamycin, an inhibitor N-linked glycosylation, and DTT, which prevents disulfide-linked folding. The cellular RNA was then isolated and blotted, followed by hybridization with ^{35}P-labeled probes to detect TMX4 and BiP mRNA. While the UPR-target BiP was upregulated upon tunicamycin and DTT treatment, TMX4 mRNA levels stayed the same as in untreated cells as well as in cells that were heat-shocked, a form of stress that does not induce the UPR.

Taken together, these results indicate that the TMX4 gene is widely expressed in human tissues and is not a UPR-target.

A



B

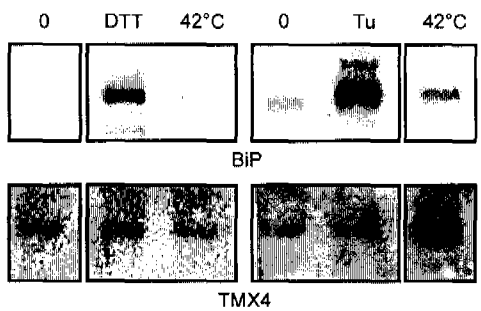


FIG. 6. **Transcriptional analysis of TMX4.** A, The TMX4 mRNA is expressed in a wide array of tissues. A human tissue Northern blot was radioactively probed for the presence of TMX4 transcripts (predicted transcript length: 6.1 kb) and analyzed by autoradiography. Note a background band (*) in all tissues that migrates at the height of 28S ribosomal RNA (4.8 kb). As a control, the blot was stripped and labeled against β-actin (2.0 kb transcript length with an additional 1.8 kb isoform in muscle tissues). B, TMX4 mRNA is not upregulated by UPR. To induce ER stress, HeLa cells were treated with 5 mM DTT and 10 μg/ml tunicamycin (Tu) for 6 h. As a control for non-UPR-inducing stress, cells were heatshocked for 30 min at 42°C. Total RNA was isolated and analyzed by Northern blot against TMX4. Subsequently, the blot was stripped and hybridized against BiP that served as positive control for the UPR upregulation.

2.6 Antibody production

To study endogenous TMX4, we produced peptide antibodies against TMX4 using a segment of the cytosolic tail, EANDQGPPGEDGVTR (residues 295 to 309), as antigenic peptide. Peptide production, coupling to keyhole limpet hemocyanin and immunization of two rabbits was performed by Eurogentec. Due to the low signal both antisera gave (data not shown), one of the antisera was purified by affinity-

chromatography using a column that had before been coupled to the antigenic peptide. In spite of the purification step, the antiserum recognized overexpressed HA-TMX4 from HeLa cells on Western blots only weakly and did not detect endogenous TMX4 at all (data not shown). No further experiments were performed with this antiserum. Instead, the whole cytosolic tail of TMX4 (TMX4cyt), fused with a hexa-histidine-tag, was expressed in *E. coli*, purified over a Ni^{2+}-NTA column followed by anion exchange chromatography and handed to Gregor Fischer, Institut für Labortierkunde, University of Zurich, for immunization of two rabbits. One of the antisera was affinity purified using TMX4cyt immobilized on polyvinylidene difluoride membrane strips. Neither the unpurified, nor the purified antisera recognized endogenous TMX4 or overexpressed HA-TMX4 from HeLa cells on Western blots (data not shown).

2.7 Determination of the TMX4 redox state in vivo

Knowledge about the *in vivo* redox state of the active-site cysteines in an oxidoreductase can help elucidate the redox function performed by the enzyme. Here, we determined the oxidation state of metabolically labeled HA-TMX4 expressed in HeLa cells. The protocol used involved a two-step alkylation process of thiols; first with NEM added to living cells and then – after lysis, *in vitro* treatment with the reducing agent TCEP and immunoprecipitation – with AMS (see Materials and Methods). By this method, reduced cysteines will be modified with NEM, whereas oxidized cysteines become modified with AMS. Due to the larger size of the latter reagent compared to NEM (450 Da versus 125 Da) oxidized proteins will be separated from reduced proteins by SDS-PAGE as a result of their slower mobility. Cells treated with DTT and diamide were used to obtain control samples for the reduced and oxidized state, respectively.

As shown in Fig. 7, lane 2, we detected two bands for AMS-modified HA-TMX4. As expected, the lower band migrated like the reducing control (lane 3). Surprisingly though, the upper band did not run at the same height as the oxidizing control (lane 4), which showed an unexpected fast mobility. A control experiment showed that this not was caused specifically by diamide since another oxidizing reagent, dipyridyl disulfide, produced the same effect on TMX4 (data not shown). Moreover, the effect was specific for TMX4. This was shown by using another PDI-family member, TMX3, which like TMX4 contains a single thioredoxin-like domain

(Haugstetter et al., 2005), for comparison (Fig. 7, lower panel). In this experiment, we detected TMX3 by Western blotting in the exact same samples used in the upper panel for the HA-TMX4 experiment. At steady state (lane 6), TMX3 showed the expected distribution of about 30% oxidized, 70% reduced (Appenzeller-Herzog and Ellgaard, 2007; Haugstetter et al., 2005) and, importantly, the oxidized band ran at the height of the material in the diamide-treated controls (lanes 4 and 8) showing that diamide treatment was indeed effective.

As mentioned above, TMX4 contains five cysteines in total. To ensure that the upper band detected for HA-TMX4 in lane 2 was derived from the two cysteines in the active site, we mutated the CPSC active-site sequence of TMX4 to CPSS thereby preventing the formation of a disulfide bond. The result (Fig. 7, lanes 5-8, upper panel) showed the appearance of only a single band for HA-TMX4/CPSS at steady state (lane 6) migrating like the reduced control. We concluded that the upper band in lane 2 corresponds to oxidized HA-TMX4 and determined by densitometry this fraction to be ~70%, with ~30% in the reduced state.

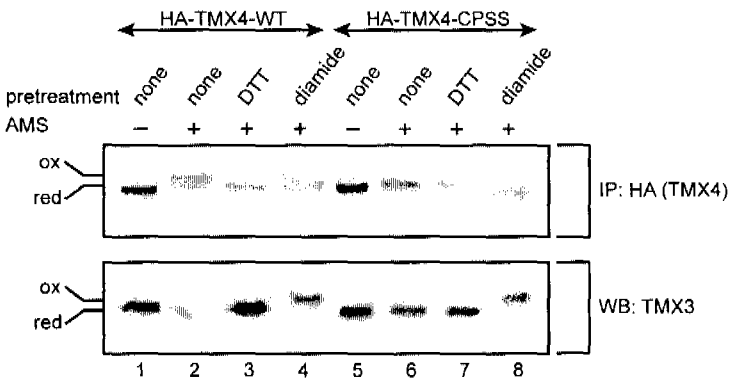


FIG. 7. **HA-TMX4 is partially oxidized *in vivo*.** HeLa cells transiently expressing wild-type HA-TMX4 (active-site sequence: CPSC; lanes 1-4) or the CPSS active-site mutant (lanes 5-8) were metabolically labeled, treated with 20 mM NEM, lysed and immunoprecipitated with an anti-HA antibody (autoradiogram shown in upper panel). Following reduction by TCEP, the lysates were mock treated or treated with AMS and then separated by SDS-PAGE. Cells were pretreated with 10 mM DTT and 5 mM diamide to obtain controls for the fully reduced and oxidized forms, respectively. Lower panel: Western blot against TMX3 performed on the same lysate as the TMX4 immunoprecipitation. The positions of the oxidized and reduced forms of both proteins are indicated. The atypical mobility of TMX3 in lane 2 is very likely caused by a gel artefact.

2.8 *In vitro studies of TMX4*

The results shown in the following three sections were generated by Nils Althaus during the course of his diploma thesis performed under my supervision.

For *in vitro* studies of basic enzymatic properties, redox potential measurement and structure determination of TMX4, we expressed the thioredoxin-like domain and purified it from *E. coli*. To this end, we generated three constructs, all hexa-histidine-tagged at the N-terminus for Ni^{2+}-NTA affinity purification (Fig. 8): His-TMX4-Trx, containing residues 26 to 144 of TMX4; His-TMX4-Lum comprising additionally the residues 145 to 185, predicted to be hydrophobic (Fig. 5B) and His-GB1-TMX4-Trx consisting of the TMX4 thioredoxin-like domain fused to a cleavable GB1-domain of Streptococcus spec. (Card and Gardner, 2005; Huth et al., 1997; Zhou et al., 2001).

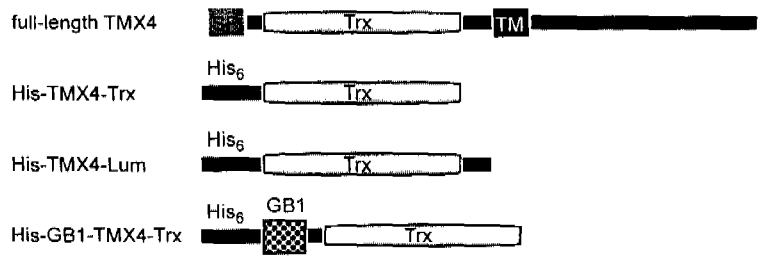


FIG. 8. **Three different His-tagged TMX4 constructs were used for expression in *E. coli*.** One comprised only the thioredoxin-like domain, the second extended over the entire predicted lumenal domain of TMX4 and the third was a cleavable fusion construct between the highly soluble GB1-domain of Streptococcus spec. and the thioredoxin-like domain of TMX4. All three products were tagged with six N-terminal histidine resides for Ni^{2+}-NTA affinity purification (His_6).

2.8.1 Expression and purification of His-TMX4-Trx

Under standard protein expression conditions of 37 °C, the small His-TMX4-Trx (136 residues, 15.5 kDa) was found in inclusion bodies (data not shown). To enhance solubility, His-TMX4-Trx was expressed at different temperatures. Expression levels at 16 °C (and even more so at 4 °C) were lower than at 25 °C, and the amount of soluble overexpressed protein could not be significantly enhanced (data not shown).

As solubility of His-TMX4-Trx could not be increased, the strategy was changed, the same construct was purified in a denatured state and refolded on a Ni^{2+}-

NTA column. After incubation for two weeks at 4 °C, however, the protein began to precipitate. Moreover, when we attempted to reproduce this method of refolding several months later, His-TMX4-Trx eluted only with 100 % EDTA-urea elution buffer, which stripped the Ni^{2+}-ions and solubilized any aggregated protein. This result implied that His-TMX4-Trx had not efficiently been refolded on the column (data not shown).

We pursued two other strategies: refolding by dialysis as well as refolding by rapid dilution. First, 6 ml of eluate from the Ni^{2+}-NTA chromatography were dialyzed into non-denaturing resuspension buffer. At 4 °C, the protein precipitated continuously over two days until - after removing the precipitate - the concentration of protein in solution reached 0.37 mg/ml. Second, 1.5 ml of eluate were rapidly diluted into a 10 fold excess of non-denaturing resuspension buffer, but also with this approach, precipitates formed until after two days, the concentration of protein in solution remained stable at 0.58 mg/ml.

For further purification by anion exchange chromatography, His-TMX4-Trx refolded by dialysis was dialysed into low salt buffer, after which the concentration was measured to be 0.32 mg/ml, slightly lower than at high salt conditions. The sample was loaded on an anion exchange column and eluted with high salt buffer (Fig. 9). The sharp peak at 22 ml in the elution profile in Fig. 9A contained pure soluble His-TMX4-Trx. The protein concentration of the fraction between 19 and 23 ml was only 0.13 mg/ml. The other fractions of the elution with high salt buffer were less pure. When denaturing high salt buffer was applied on the column, thereby solubilising any aggregated or unfolded protein still bound on the column, His-TMX4-Trx eluted in a sharp and intensive peak (Fig. 9A). This result again implied considerable precipitation on the column. On the SDS-PAGE gel, apart from the expected band at 15.5 kDa, another band at approximately twice the size was detected, possibly containing His-TMX4-Trx dimers (Fig. 9B). Already in earlier experiments we had observed such formation of bands corresponding to dimers - in spite of the reducing and denaturing conditions of sample preparation and SDS-PAGE analysis - and confirmed their identity by mass spectrometry (data not shown).

We therefore wondered whether even the seemingly soluble His-TMX4-Trx was in fact aggregated. To detect aggregations invisible by eye, an absorption

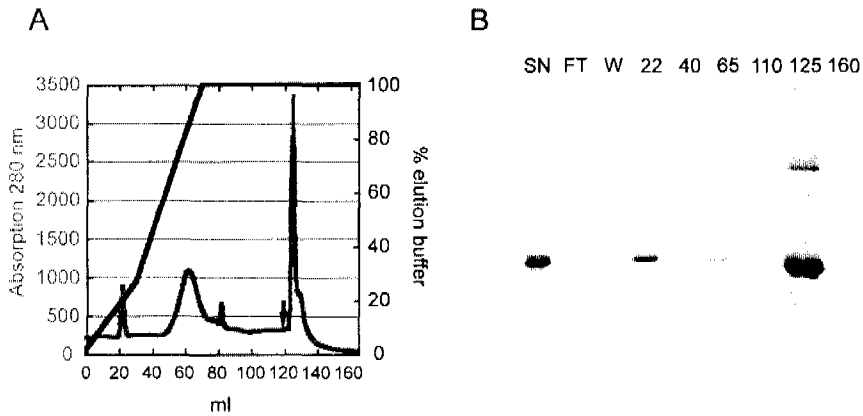


FIG. 9. **His-TMX4-Trx refolded by dialysis was highly purified by anion exchange chromatography.** A, Elution profile of His-TMX4-Trx anion exchange chromatography. The arrow indicates a change from non-denaturing high-salt buffer to denaturing high-salt buffer containing urea. The blurring of the bands is a gel artifact and was not seen on other gels. B, Protein gel corresponding to elution profile. Load (L), flowthrough (FT), wash (H). Numbers correspond to fractions at the indicated positions in the elution profile in ml.

spectrum was recorded of the pooled fractions of the anion exchange chromatography. The spectrum had a minimum at approximately 250 nm, a maximum at approximately 280 nm and a flat baseline at wavelengths above 310 nm, two properties characteristic for soluble proteins (data not shown).

We then went on to assess the secondary structure content of His-TMX4-Trx by circular dichroism (CD) measurements (Fig. 10B and C). With a local minimum at 208 nm and one at 222 nm, the far-UV CD spectrum at 4°C corresponded to the characteristic shape for a protein with significant α-helical content. However, the molar ellipticity between 210 and 220 nm did not reach values as low as expected for a fully folded thioredoxin-like domain, which contains α-helices as well as β-strands (e.g.(Frickel et al., 2004; Haugstetter et al., 2005; Philipps and Glockshuber, 2002)).

In order to determine the temperature of unfolding for His-TMX4-Trx, the molar ellipticity was followed at 220 nm upon heating of the protein solution from 4 °C to 85 °C (temperature transition curve, Fig. 10C). The polypeptide is partially unfolded at lower temperatures before it unfolds between 25 °C and 37 ° (black line). The unfolding was irreversible (red line). It is also notable that after the first temperature transition, precipitate was visible by eye in the cuvette.

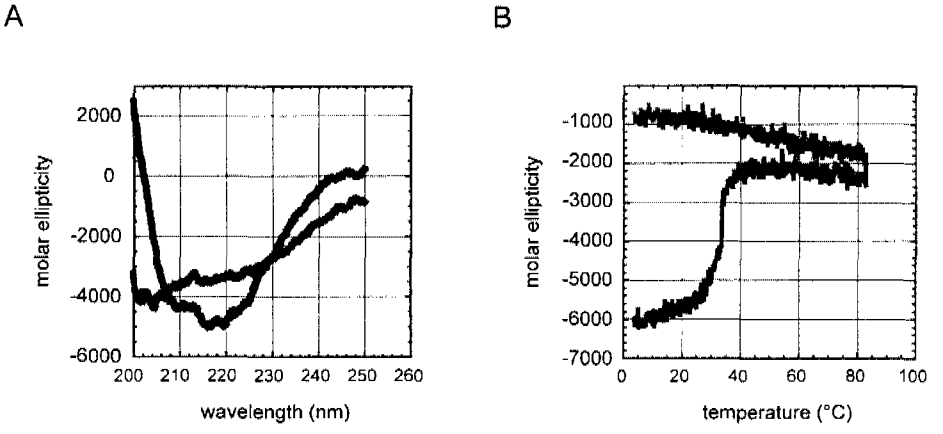


FIG. 10. **His-TMX4-Trx refolded by dialysis is unstable at room temperature.** A, Circular dichroism spectrum of purified His-TMX4-Trx. Spectrum recorded at 4 °C (black line) and 95 °C (red line). B, Purified His-TMX4-Trx unfolds at room temperature and even below. Measurement of the molar ellipticity of purified His-TMX4-Trx at 220 nm at temperatures between 4 °C and 85 °C. first measurement (black line) and second measurement after incubation of the sample for 1 h at 4°C (red line).

Based on these results it was clear that the refolded His-TMX4-Trx purified so far was not suitable for *in vitro* experiments. In order to test whether refolding of His-TMX4-Trx was more effective under other conditions, a solubility screen was performed on denatured His-TMX4-Trx after elution from the Ni^{2+}-NTA column with EDTA-urea elution buffer (Appendix, table 2). The eluate was refolded by dialysis overnight into different buffers at different pH and salt concentrations and then examined for precipitation under a light microscope. The screen showed that slightly basic high-salt Tris-HCl buffers were most favorable for the solubility of His-TMX4-Trx.

A second solubility screen was performed on the same protein solution in order to test different reagents potentially counteracting protein aggregation (Appendix, table 3). The results showed that the only reagent conferring some degree of solubility to His-TMX4-Trx was L-arginine. After removal of L-arginine by dialysis, however, prominent precipitation was observed again. Changing the redox conditions by omitting DTT did not promote solubility either. Both screens confirmed the initial choice of buffer conditions used for refolding, but they also showed that no other conditions could be found that would increase the solubility of His-TMX4-Trx.

2.8.2 Expression and purification of His-TMX4-Lum and His-GB1-TMX4-Trx

The His-TMX4-Trx construct proved to yield a very aggregation-prone protein. Given that the linker region between the thioredoxin-like domain and transmembrane helix (residues 145 to 185) was predicted to be hydrophobic (Fig. 5B), we considered that this linker could cover a potential hydrophobic patch on the thioredoxin-like domain. We therefore expressed the whole lumenal domain, His-TMX4-Lum (178 residues, 20.1 kDa). Like the thioredoxin-like domain alone, however, His-TMX4-Lum was insoluble when expressed at different temperatures and was inefficiently refolded on a Ni^{2+}-NTA column (data not shown). When the Ni^{2+}-NTA fractions were dialyzed into low salt buffer before further purification, His-TMX4-Lum precipitated even stronger than His-TMX4-Trx. We concluded that His-TMX4-Lum was even more difficult to work with than His-TMX4-Trx and did not conduct further experiments with this construct.

The third construct, His-GB1-TMX4-Trx, contained the highly soluble GB1 domain (6 kDa, 54 residues) that is used to hold fused proteins in solution (fig. 8) (Card and Gardner, 2005; Huth et al., 1997; Zhou et al., 2001). Indeed, expression at 37 °C and even more so at 25 °C resulted partially in soluble protein (data not shown) that could be purified by Ni^{2+}-NTA chromatography. The fractions, however, contained only 0.16 mg/ml soluble protein. 20 hours after proteolytic cleavage of GB1 from the thioredoxin-like domain via a built-in thrombin cleavage-site, precipitation was visible by eye. Since His-TMX4-Trx was soluble to a concentration of 0.37 mg/ml and thus more soluble than the cleavage product of His-GB1-TMX4-Trx in the same buffer, we decided not to do any further work with the His-GB1-TMX4-Trx construct.

2.9 *Co-Immunoprecipitation of TMX4 with binding partners*

Aside of studying the *in vivo* redox state and *in vitro* redox properties, identifying substrate proteins of TMX4 might contribute to elucidating the function of the enzyme. Thiol-disulfide oxidoreductases and their substrates form mixed disulfides which are of such transient nature that they are difficult to trap. Since intramolecular disulfide-bond formation between the two active-site cysteines leads to substrate

release, the interaction between enzyme and substrate can be stabilized by replacing the C-terminal active-site cysteine by a serine. HA-TMX4 wild-type and the CPSS active site mutant were expressed in HeLa cells that were then metabolically labeled with ^{35}S cysteine and methionine. The following steps were performed under conditions that prevented post-lysis thiol-disulfide exchange reactions. HA-TMX4 wild-type and HA-TMX4-CPSS were immunoprecipitated in the presence of NEM and analyzed by non-reducing SDS-PAGE. The autoradiograms did not reveal any high molecular weight bands that could correspond to a complex between TMX4 and intracellular binding partners (data not shown).

To enhance the chances of isolating a complex between TMX4 and substrates, transfected HeLa cells were metabolically labeled, before they were incubated with the membrane-permeable, homobifuncional compounds BMH and DSS. These chemicals react with free thiols and amino groups and can crosslink proteins. Isolation, analysis and visualization were carried out as described above. Yet, no bands representing TMX4 and a protein adduct were observed by autoradiography (data not shown). The lack of success might be due to a detection problem caused by the limited amount of cellular material available when transiently expressing proteins.

2.10 An Arg-based signal is sufficient for ER localization

2.10.1 A non-classical motif is involved in ER targeting of TMX4s

Although TMX4 resides in the ER, it lacks a classical K(X)KXX-type ER localization signal that targets many type I membrane proteins to the ER. In our multiple sequence alignments of TMX4 we noticed a LRQR motif conserved among species and found close to the C-terminus in the otherwise relatively less well conserved cytosolic tail (Fig. 1B).

To test the possible involvement of the LRQR sequence in the ER localization of TMX4, we produced two HA-tagged mutants – a truncation variant lacking the C-terminal 121 residues (Δtail) and a mutant where the two arginines of the RQR sequence had been substituted by lysines (KQK) (Fig. 11A). The latter rather conservative mutation has been shown to efficiently abolish the function of the RXR sequence in Kir6.2, a subunit of K_{ATP} channels (Zerangue et al., 1999). After

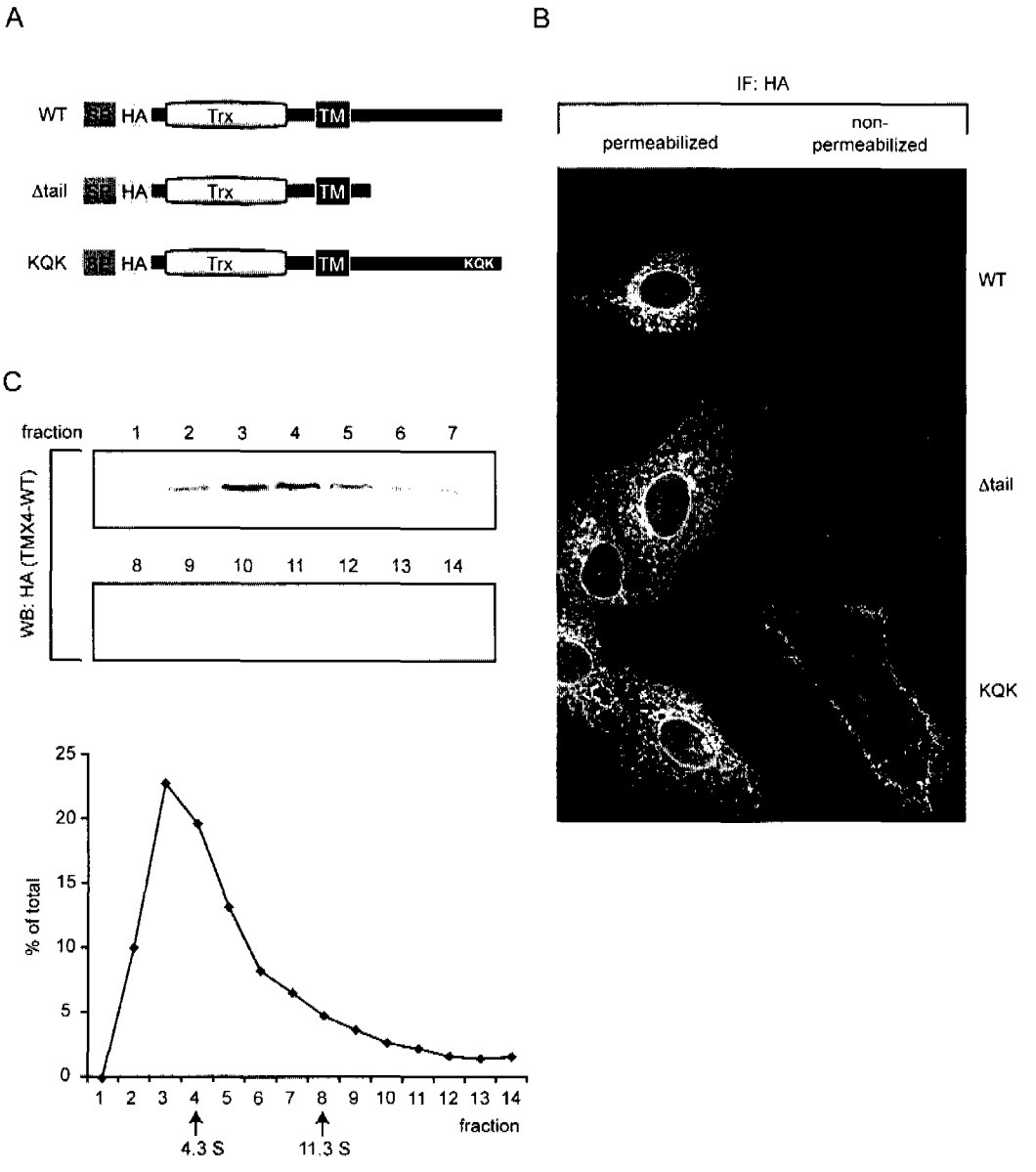


FIG. 11. **The cytoplasmic RQR motif contributes to ER localization of TMX4.**
A, schematic representation of the HA-TMX4 mutants variants used for localization studies: the full-length wild-type HA-TMX4 (WT), HA-TMX4 lacking the cytoplasmic region (Δtail), the HA-TMX4 mutant with the two arginines in the native RQR sequence substituted by lysines (KQK). Further abbreviations as in Fig. 1A. B, mutants lacking the RQR motif partially escape the ER and reach the plasma membrane. Vero cells were transfected with the three TMX4 constructs and stained with the anti-HA antibody after fixation and permeabilization (left). For detection of protein on the cell surface, cells were incubated with anti-HA antibody before fixation (right). IF: immunofluorescence. C, HA-TMX4 behaves as a monomer in velocity centrifugation. Postnuclear supernatants of HA-TMX4-transfected HeLa cells were supplemented with 0.1% Triton X-100, followed by velocity centrifugation on 5-25% continuous sucrose gradients at 42000 rpm for 16 h at 4°C. After harvesting the gradient from top to bottom, the fractions were analyzed by SDS-PAGE and Western blotting (upper panel). The HA-TMX4 signal was quantified by densitometry and plotted as the percentage of total (lower panel). As markers, bovine serum albumin (4.3 S; 66 kDa) and β-amylase (11.3 S; 200 kDa) were analyzed on parallel gradients and the peak fraction was determined by a Bradford assay.

transfecting Vero cells, we analyzed the cellular localization of the wild-type protein and the two mutants by immunofluorescence microscopy (Fig. 11B). We found that, unlike the wild-type protein, HA-TMX4 lacking the cytosolic tail partially escaped the ER and reached the plasma membrane, as visualized in non-permeabilized cells. The KQK mutant showed the same staining pattern, being detected in the ER but also on the cell surface.

While some of the mutant TMX4 molecules trafficked to the cell surface, the majority of the all three HA-TMX4 variants localized to the ER. To ascertain that this retention was not caused by aggregation as a result of over-expression we used sucrose gradient velocity centrifugation. By this technique, aggregates sediment more slowly than monomeric or oligomeric proteins. HeLa cells over-expressing HA-TMX4 were solubilized in Triton X-100 and the lysates loaded onto a 5-25% continous sucrose gradient. After centrifugation, manually collected fractions were separated by SDS-PAGE and analyzed by Western blotting. For comparison, we ran in parallel samples of bovine serine albumin (BSA, a monomeric protein of 66 kDa that sediments at 4.3 S) and β-amylase (a tetrameric protein of 200 kDa for the whole complex that sediments at 11.3 S) (Johansson et al., 2001) on the same type of gradients and determined the protein content of the fractions by a Bradford assay. Quantification of the HA-TMX4 signals showed a discrete peak of the protein in fraction 3, ahead of BSA that peaked in fraction 4 (Fig. 11C). The distribution of HA-TMX4 was relatively broad so that the presence of a small amount of lower oligomeric complexes, for instance containing HA-TMX4 and a substrate, could not be ruled out. Clearly though, no evidence was found for high-molecular weight complexes. The experiment showed that HA-TMX4 does not form aggregates in the cell, and therefore that ER retention of HA-TMX4 is not an artifact of protein over-expression. We concluded that the RQR sequence in the C-terminal region contributes to ER localization of HA-TMX4.

In addition to Δtail and HA-TMX4-KQK, we made a construct consisting of only the ER-lumenal part of HA-TMX4, which we named HA-TMX4-lum. It was expected to be secreted and immunoprecipitated from the cell culture medium. As it could not be detected in the medium, we wondered whether it was not secreted but instead degraded intracellularly. To test this possibility, we transfected HeLa cells that

were then metabolically labeled with 35S-methionine/cysteine and chased for up to 3 h in absence or presence of the proteasome inhibitor MG132 (Fig. 12A). By densitometry, we quantified the band signals and estimated the half life of HA-TMX4-lum to be approximately 40 min. When the cells were treated with MG132, the protein was clearly stabilized (Fig. 12B). These results suggest that HA-TMX4-lum is rapidly degraded by ERAD and might therefore fail to be secreted.

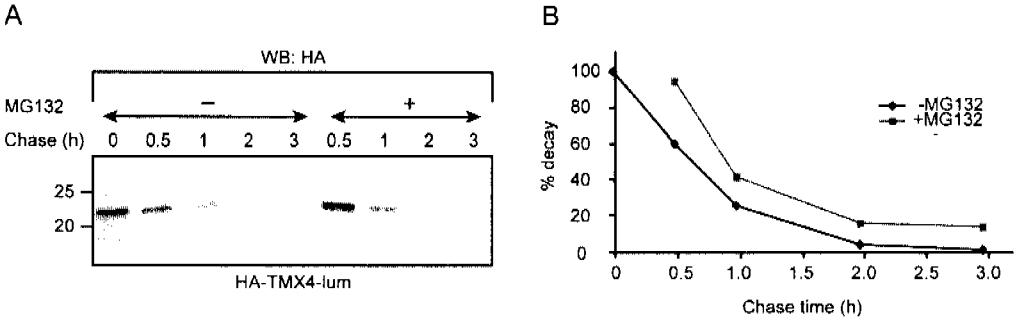


FIG. 12. **Degradation of HA-TMX4-lum is delayed by proteasome inhibitor MG132.** A, Transfected HeLa cells were metabolically labeled with ^{35}S for 5 min and chased for the indicated periods of time in absence or presence of 10 μM MG132. Cells were then lysed, subjected to immunoprecipitation and analyzed by SDS-PAGE and autoradiography. B, Densitometric quantification of the autoradiogram in A. The percentage of protein at chase time 0 was set to 100%, and the percentages of remaining protein was plotted against the chase time.

2.10.2 Surface exposed TMX4-KQK remains Endo H-sensitive

As mentioned above, the Endo H-sensitive high-mannose *N*-glycans found on ER proteins are usually modified to complex, Endo H-resistant glycans in the medial Golgi. We therefore expected that HA-TMX4-KQK would become Endo H-resistant while being exported to the plasma membrane.

To follow the anterograde transport of HA-TMX4-KQK, we expressed the protein in HeLa cells and performed a pulse/chase assay in combination with Endo H-digestion (Fig. 13A). In contrast to our expectations, HA-TMX4-KQK is deglycosylated by Endo H even after 2 hours of chase. ER retention of the mutant was unlikely an explanation for the persistent Endo H-sensitivity because we knew from our immunofluorescence experiments (Fig. 11B) that HA-TMX4-KQK does reach the cell surface. As, however, this is only a minor fraction of the protein, it is possible that surface-exposed HA-TMX4-KQK is under the detection limit.

Occasionally, *N*-glycans do not receive the modifications that are responsible for Endo H-resistance while passing the Golgi apparatus. To test whether this is the case for HA-TMX4-KQK, we crosslinked biotin to surface-exposed proteins and separated them from intracellular material by means of streptavidin beads. The isolated cell surface proteins and total cell extract were then mock-treated or digested by Endo H before separation by SDS-PAGE and Western blotting (Fig. 13B). We found that plasma membrane-localized HA-TMX4-KQK remains Endo H-sensitive, rendering the pulse/chase-Endo H approach ineffective to study HA-TMX4-KQK trafficking to the cell surface.

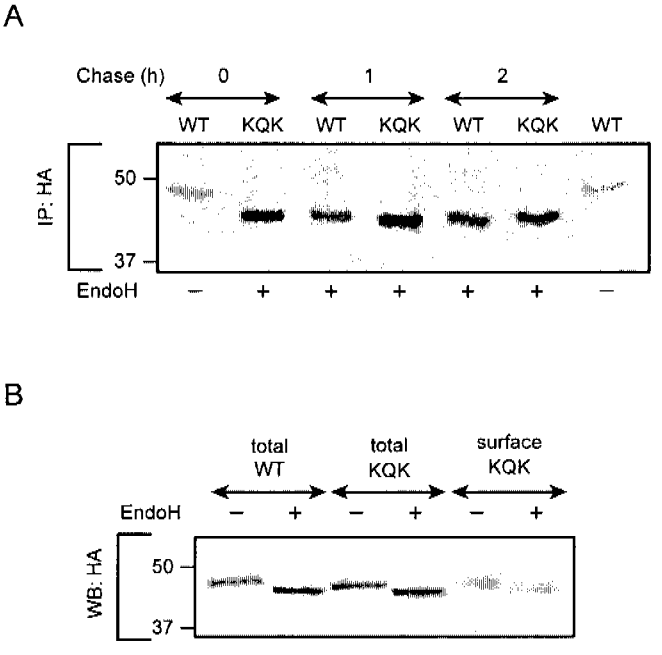


FIG. 13. **Plasma membrane exposed HA-TMX4-KQK remains Endo H-sensitive.**
A, After HeLa cells were transfected with HA-TMX4 and HA-TMX4-KQK, respectively, they were metabolically labeled with ^{35}S for 5 min and chased for the given periods of time. HA-TMX4-WT and –KQK were immunoisolated before they were mock treated or treated with Endo H. The digest mixtures were analyzed by SDS-PAGE and autoradiography. B, HeLa cells were transfected as in A and either lysed directly to yield total lysate or first subjected to crosslinking of cell surface proteins to Biotin and subsequent isolation with Strep-Tactin beads. Total lysate (total) and isolated cell surface protein (surface) were mock-treated or treated with Endo H before they were separated by SDS-PAGE and visualized by Western blot.

2.10.3 ER localization conferred by the TMX4 cytosolic tail depends on the RQR sequence

To further characterize the importance of the TMX4 LRQR sequence in ER retention, we created three chimeric proteins (Fig. 14A). All were HA-tagged chimeras between the luminal region of the plasma membrane protein CD4 (Maddon et al., 1985) and the cytosolic tail of TMX4. Two fusion proteins contained the wild-type sequence for the cytosolic tail of TMX4 but differed with respect to the origin of the transmembrane domain (TMX4 or CD4), whereas the last chimera contained the CD4 luminal and transmembrane domain fused to the KQK mutant of the TMX4 cytosolic tail.

The chimeras and wildtype HA-CD4 were transfected into HeLa cells, which were then analyzed by confocal immunofluorescence microscopy. As expected, we found HA-CD4 strongly staining the plasma membrane. In contrast, the two CD4 chimeras fused to a wild-type TMX4 tail sequence were completely retained in the cell and did not reach the cell surface (Fig. 14B). The nature of the transmembrane domain – whether from CD4 or TMX4 – did apparently not influence ER localization. Interestingly, the chimeric protein of HA-CD4 fused to the cytosolic tail of TMX4 containing the KQK mutation was clearly detected on the plasma membrane and showed a distribution similar to wildtype HA-CD4.

Having established that simply mutating the RQR sequence in the cytosolic tail of TMX4 to KQK was enough to redirect the chimeric protein to the plasma membrane, we wanted to investigate in more detail the trafficking of the various chimeras from the ER to the cell surface. This was done using a pulse-chase approach in combination with immunoprecipitation and Endo H digestion. The four HA-tagged CD4-fusion proteins were expressed in HeLa cells and metabolically labeled with ^{35}S cysteine and methionine. Following chase-times of up to 12 h, the proteins were isolated by immunoprecipitation, before they were either digested with Endo H or mock-treated. Upon SDS-PAGE, the gels were then dried and the proteins detected by autoradiography (Fig. 15A).

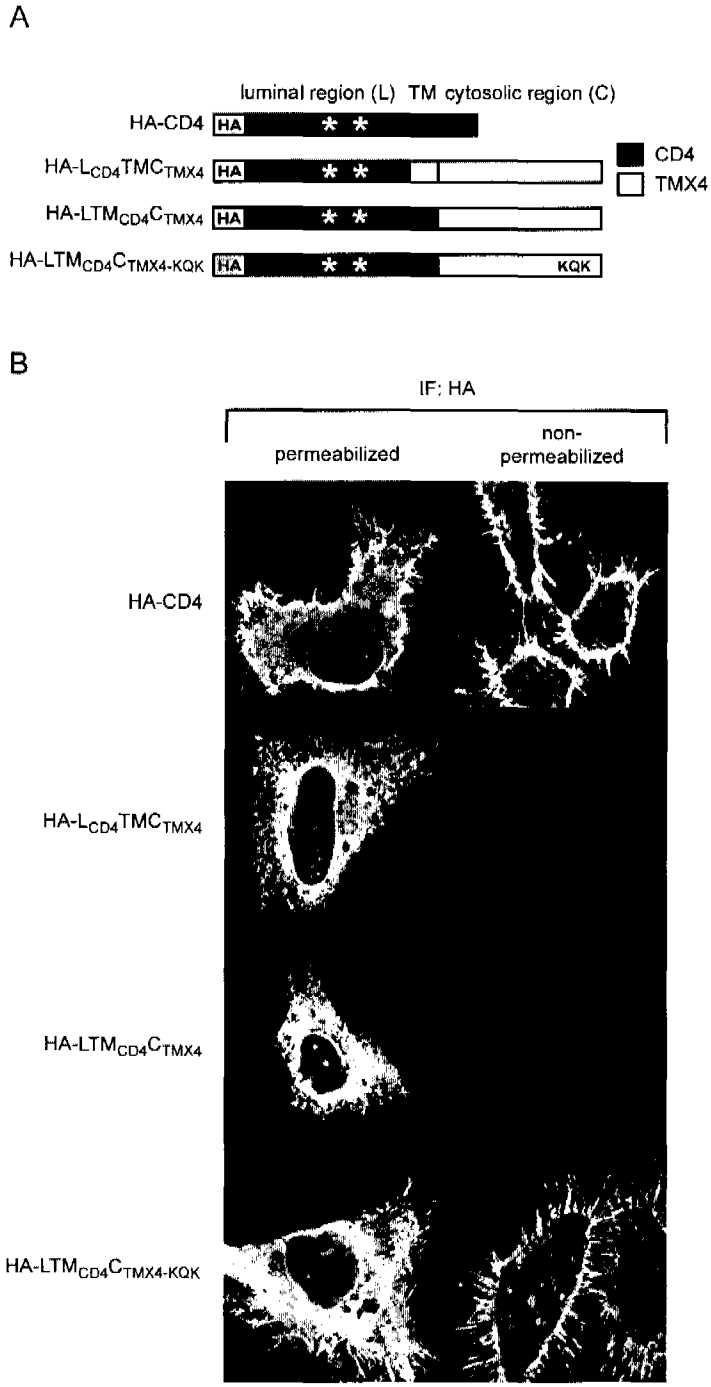


FIG. 14. **The TMX4 RQR sequence prevents the cell surface expression of the plasma membrane protein CD4.** A, schematic overview of HA-tagged CD4 and CD4/TMX4 chimeras. All possess the ER-luminal region of CD4 (L_{CD4}) but differ in transmembrane domain (TM) and cytosolic tail (C), which are from CD4, TMX4 or TMX4-KQK as indicated. The N-glycosylation sites of CD4 are marked by an asterisk (*). B, CD4 fusion proteins carrying the cytosolic region of WT-TMX4 do not reach the cell surface, whereas the cytosolic tail of TMX4-KQK is not sufficient to retain CD4 intracellularly. HeLa cells were transfected with the HA-CD4/TMX4 constructs, fixed and permeabilized or not before staining with an anti-HA antibody to detect intracellular and cell surface expression, respectively, of the HA-CD4 chimeras.

As indicated in Fig. 14A, the luminal region of CD4 contains two *N*-glycans. When treating lysates with Endo H, the results showed that only one of the glycans became Endo H-resistant as CD4 trafficked to the plasma membrane, whereas the other remained Endo H-sensitive. This is most clearly seen in the upper panel of Fig. 15A, where after chase times of 0.5, 5 and 12 h two bands appeared upon Endo H-digestion – the lower migrated as the completely sensitive protein (compare to the 0 h chase time, + Endo H) whereas the band that appeared after 0.5 h chase time ran in between the glycosylated and the fully deglycosylated forms. This partially Endo H-resistant form (labeled R in Fig. 15A) was used as a measure of CD4 on the plasma membrane.

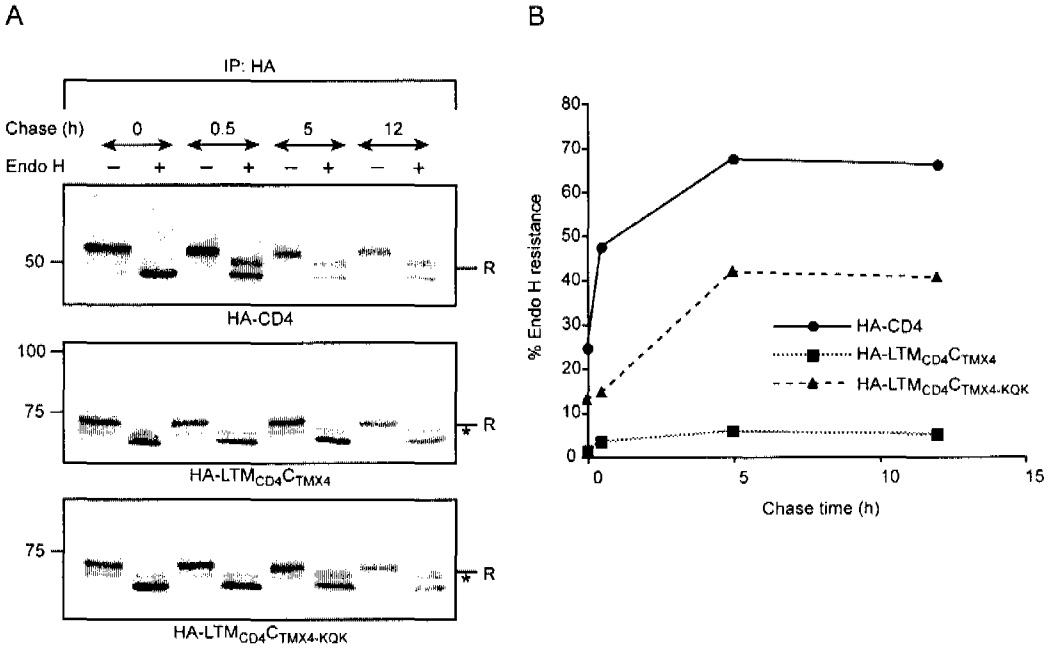


FIG. 15. **The TMX4 RQR motif is required for ER localization of the plasma membrane protein CD4.** A, HeLa cells were transfected with HA-CD4 and chimeric HA-CD4/TMX4 constructs as indicated, metabolically labeled with ^{35}S for 5 min and chased for the given periods of time. After immunoprecipitation, the HA-CD4/TMX4 variants were mock treated or treated with Endo H before the digest mixtures were separated by SDS-PAGE and visualized by autoradiography. Note that the band that represents the Endo H-resistant form (labeled 'R') migrates in between the undigested and completely sensitive forms. A background band that migrates slightly below the Endo H-resistant forms of HA-LTM$_{CD4}$-C$_{TMX4}$ and HA-LTM$_{CD4}$-C$_{TMX4-KQK}$ is labeled with an asterisk (*). B, Densitometric quantification of the autoradiograms in A. The percentage of Endo H-resistant protein was plotted against the chase time.

For the two fusion proteins with TMX4 (middle and lower panel, Fig. 15A) the R band partially overlapped with a background band (labeled *). To correct for this overlap, we first quantified the two bands together and then subtracted the intensity of

the background band from the corresponding mock-treated lane. While this method did perhaps not completely prevent minor quantification errors, the results from four independent experiments were clear. We found that the Endo H-resistance reached 65% for wild-type HA-CD4, whereas the construct with the wild-type TMX4 tail, HA-LTM$_{CD4}$C$_{TMX4}$, remained completely sensitive (Fig. 15B). Full Endo H-sensitivity was also observed for the chimera that contained the transmembrane domain of TMX4 as well (data not shown). Importantly, when the RQR sequence was mutated to KQK, the Endo H-resistance of the fusion protein was restored to 40%. These results indicated that the cytosolic tail of TMX4 is sufficient for ER targeting and that the RQR sequence is required for this ability.

3 Discussion

3.1 Active site and potential redox function of TMX4

The number of published proteins belonging to the human PDI family has increased significantly in recent years. A striking feature of TMX4 and its paralog TMX are their unusual active motives. By far most of the TMX4 orthologs, including human, contain the active-site sequence CPSC, whereas all TMX proteins have a CPAC motif. Common active sites found in efficient oxidases are CXHC, for example CPHC in the bacterial DsbA, the most oxidizing thiol-disulfide oxidoreductase, and CGHC in mammalian PDI. For the PDI family members, the CGHC sequence is the most common with 16 occurrences among 25 domains with a CXXC sequence. In contrast, efficient reductases often use CXPC like CGPC in bacterial thioredoxin itself (Ellgaard and Ruddock, 2005).

To understand the role of the two X residues between the active site cysteines, it is important to have in mind that the mechanism of disulfide exchange reactions begins when the the N-terminal cysteine (cys_N) of the CXXC motif attacks a disulfide bond nucleophilically. The cys_N exhibits a very low pKa which is mainly caused by hydrogen bonds (Forman-Kay et al., 1991; Jeng et al., 1994) and interaction with the partial positive charge of the dipole of the helix, at which N-terminus the active site is positioned (Hol, 1985; Sali et al., 1988; Serrano and Fersht, 1989; Walter et al., 1995). As expected, studies with DsbA active site mutants showed a strong correlation between the redox potential and the acidity of cys_N (Grauschopf et al., 1995). The lower the pKa, the more oxidizing the enzyme. The pKa of cys_N in turn is modulated by the XX dipeptide of the active site.

At first glance, the TMX and TMX4 active sites seem to resemble rather the oxidizing motif CPHC (DsbA) than the reducing CGPC (thioredoxin). The common proline at the first position of the XX dipeptide in the active site has been found to be the most adequate residue for the stabilization of the cys_N thiolate by the helix dipol (Kortemme and Creighton, 1995; Quan et al., 2007). Nevertheless, it would be uncareful to assume TMX and TMX4 must function as oxidases. The histidine in the CPHC active site of wildtype DsbA has been suggested to stabilize the thiolate anion of cys_N, thus lowering the pKa and contributing considerably to DsbA's oxidizing power. Replacing the histidine by other residues weakens the oxidizing activity of the

DsbA mutant, down to a mutant (CPPC) as reducing as thioredoxin (Grauschopf et al., 1995). Hence if TMX and TMX4 are oxidases, they are likely to be less oxidizing than DsbA, with TMX eventually being a little less oxidizing than TMX4 which has a serine in the CPSC active site that could stabilize the thiolate anion better than the alanine in CPAC.

Interestingly, an exception to the strong correlation between cys_N pKa and redox potential of DsbA mutants was a CPGC variant which was more oxidizing than expected, given the relatively high cys_N pKa of 4.85 (Huber-Wunderlich and Glockshuber, 1998). In contrast to DsbA, the correlation between the redox potential and the pKa of cys_N deviates for all thioredoxin mutants, indicating that the pKa of cys_N is not the only factor that determines the redox potential. Additional factors include the pH, the titration of residues other than thiols and noncovalent interactions between the CXXC motif and surrounding residues. The determination of the redox potential is thus complex. It can prove difficult to predict the precise redox potential based alone on how the XX dipeptide might influence the pKa of cys_N.

The best estimate for the TMX4 active site is expected to be provided by two studies where various XX mutants of DsbA (Bessette et al., 2001; Quan et al., 2007) were tested for redox potential or activity *in vivo* (they include only CPSC but not CPAC). In spite of the above-mentioned considerations about the proline, alanine and serine, CPSC behaves more reducing than PDI in the two studies. This data stands in opposition to the steady state redox state of TMX4 that we measured *in vivo*, showing that 70% of TMX4 have an oxidized active site. Although it is usually assumed that a thiol-disulfide oxidoreductase is mainly present in a state ready to react with a substrate, i. e. for example an oxidase is mainly present in oxidized form, it is conceivable that kinetic restrictions influence the steady state of TMX4. On the other hand, the protein context matters. It is therefore likewise imaginable that the CPSC active site shows a different redox potential in DsbA than in its original environment, TMX4. Moreover, due to the lack of anti-TMX4 antibody, the *in vivo* redox state of TMX4 was determined using overexpressed protein. We therefore cannot exclude that the abundance of HA-TMX4 exhausted the molecules that control its redox state, so that the measured redox state deviates from the redox state of the endogenous protein.

To obtain clues about the enzymatic reaction a thiol-oxidoreductase might catalyze, determining the redox potential is a frequently used experiment. This can be done with the full-length protein or restricted to the usually readily soluble

thioredoxin-like domains. Further, the protein of interest is often tested in *in vitro* assays for oxidase, reductase and isomerase activity. For TMX, redox activity has been demonstrated *in vitro* (Matsuo et al., 2001; Matsuo et al., 2004) but it is unclear how the protein compares in such assays to well-characterized family members such as PDI and ERp57.

3.2 The low solubility of the TMX4 thioredoxin-like domain

To investigate TMX4 redox activity, we expressed the thioredoxin-like domain of TMX4 in *E. coli* but the low solubility and circular dichroism measurements indicated that the protein was probably not fully folded (Fig. 10). Efficient folding was not promoted either by fusing the thioredoxin-like domain to the highly soluble GB1, a protein G domain of *Streptococcus spec.* GB1 kept TMX4-Trx in solution but when the domain was removed, TMX4-Trx precipitated again. Curiously, the linker region between the thioredoxin-like and the transmembrane domain is hydrophobic, as shown in the TMHMM plot in figure 5B. We therefore reasoned the linker could fold back on the thioredoxin-like domain to cover a potential, otherwise exposed, hydrophobic patch that might be responsible for aggregation. However, this construct aggregated even more heavily.

That the solubility of the human thioredoxin-like domain depends on the a stable binding partner that is absent in *E.coli* cannot be completely ruled out, even though full-length TMX4 sedimented as a monomer on a sucrose gradient (Fig. 11C). Alternatively - as TMX 4 is a glycoprotein with a single *N*-glycan on the thioredoxin-like domain - efficient folding might require interaction with lectin chaperones, mediated via the glycan that TMX4 lacks in *E. coli*. Our observation that the luminal domain of HA-TMX4 is degraded in HeLa cells (Fig. 12) also raises the possibility that the membrane association plays a crucial role for correct folding of the thioredoxin-like domain. The NMR structure of the thioredoxin-like domain of TMX (Tochio et al., 2005) (Fig. 3) reveals that a β-strand formed by three residues at the C-terminus of the structure is part of the domain. The TMX4-trx protein does not comprise the corresponding three positions. Considering the successful production of the very similar TMX-trx domain, it is possible that the domain boundaries chosen for expression of TMX4-trx were inadequate.

3.3 Differences between the sequences of TMX4 and its paralog TMX

TMX and TMX4 are paralogous to each other because they arose from a duplication event as confirmed by phylogenetic analysis. The *C. elegans* protein DPY-11 diverged earlier from a TMX/TMX4 ancestor. DPY-11 is expressed exclusively in the hypodermis and is required for body and sensory organ morphogenesis. Curiously, DPY-11 does not harbor the RQR motif but a KKTK sequence at the C-terminus indicating a different ER localization mechanism than TMX4 and TMX. Also, human TMX could not rescue a *dpy-11* mutant in *C. elegans* (Ko and Chow, 2002).

By using a multiple sequence alignment of all TMX and TMX4 orthologues available, we identified six residues that changed along the evolutionary history and thus define TMX and TMX4, respectively. Four of the six residues are in the thioredoxin-like domain which allowed us to study their position in the NMR structure of the TMX thioredoxin-like domain. Surprisingly, none of the mutations clearly suggested to have structural or functional consequences. Although highly speculative, TMX4 lacking the arginine (Arg 118 in TMX, Gly126 in TMX4) that is well conserved in many PDI-family members might indicate that TMX4 is not a very efficient enzyme. Yet, an arginine conserved in TMX4 (Arg123 in TMX4, Gln115 in TMX) in the same loop might have taken over the function of the lost arginine, rendering TMX and TMX4 equally efficient. For the two remaining of the six defining residues, it was impossible to deduce the impact of the mutations because they are located in the linker region, for which no structural data was available. Some differences between TMX and TMX4 important for e.g. function, regulation, localization, and interaction partners are therefore likely to be present in the cytoplasmic tails.

3.4 A non-classical localization motif targets TMX4 to the ER

Here, we identified an Arg-based sequence motif, LRQR, conserved in the C-terminal cytosolic tail of TMX4 throughout species that functioned as an ER-localization signal. In HA-TMX4 we found that substitution of the two arginines by lysines resulted in surface expression of the protein (Fig. 11). The finding that the two arginines of the LRQR sequence were required to confer ER localization of chimeras

with the reporter protein CD4 that normally localizes to the plasma membrane further supported that the motif constitutes a bona fide ER-localization signal. The observation that a significant fraction of the HA-TMX4-RQR→KQK mutant still remained in the ER can likely be explained by association of TMX4 with substrates or binding partners. Alternatively, the lack of a linear or conformational ER exit signal might render forward transport inefficient. It is therefore possible that – even if HA-TMX4-RQR→KQK would become Endo H-resistant in the medial-Golgi - chasing the KQK mutant for two hours was too short to follow trafficking from ER to the plasma membrane (Fig. 13).

In the pulse-chase experiment with CD4 chimeras (Fig. 15) we observed, as previously seen by others (Itin et al., 1995), that wild-type CD4 failed to achieve 100% Endo H resistance even after 12 hours chase. Given that these experiments were performed in HeLa cells that do not express CD4, the relatively slow and inefficient ER export could well reflect the lack of an endogenous system to ensure effective delivery of the protein to the plasma membrane.

Like TMX4, TMX harbors a conserved Arg-based sequence motif (VRQR) that we speculate also serves as a localization signal. It is interesting to consider why TMX4 and TMX seem to use an Arg-based signal and not the classical C-terminal K(X)KXX motif employed by most type I transmembrane proteins of the ER (Teasdale and Jackson, 1996). Several issues suggest that in order to mediate ER localization, the two motifs engage a different mechanism, different machineries, different receptors that link them to the machineries or at least different receptor binding pockets (Michelsen et al., 2005; Nufer and Hauri, 2003; Zerangue et al., 2001). First, it is important to point out that lysines and arginines in these two signals are not interchangeable. Second, the function of a di-lysine motif relies on strict spacing with respect to the C terminus (Nilsson et al., 1989), whereas it turned out that for Arg-based signals, the distance to the membrane matters. Shikano and Li (Shikano and Li, 2003) even determined two different, only little overlapping zones where di-lysine and Arg motifs, respectively, are fully active: Arg signals function 16 – 46 Å away from the membrane whereas the K(X)KXX motif remains closer to the membrane. Third, it is well established that COPI-coated vesicles carry out retrograde traffic of proteins bearing di-lysine (Cosson and Letourneur, 1994; Jackson et al.,

1993; Schroder-Kohne et al., 1998). In contrast, the role of COPI in Arg-based ER localization is based on *in vitro* binding data and at present rather speculative.

Whether Arg signals mediate ER localization via ER retention or ER retrieval is therefore a much debated question (Michelsen et al., 2005; Nufer and Hauri, 2003; Zerangue et al., 2001). In principle it would be possible to detect HA-TMX4 in the ERGIC if it was retrieved to the ER. However, from the absence of HA-TMX4 in the ERGIC we cannot conclude that HA-TMX4 is retained in the ER because in this experiment the protein might simply be under the detection limit (Fig. 4B). The inefficient export of HA-TMX4-KQK (Fig. 11) is in line with this idea. It might be informative to investigate whether the ER-localized chimera HA-LTM_{CD4}-C_{TMX4} can be found in the ERGIC. As HA-CD4 shows a much stronger cell surface signal than HA-TMX4-KQK, it is likely that HA-LTM_{CD4}-C_{TMX4} can be visualized in the ERGIC if the chimera is targeted to the ER by a retrieval mechanism.

3.5 TMX4 sorting by a potential interaction with PACS-2

In addition to the RQR motif the tails of TMX4 in particular, but also TMX, contain a long stretch of negatively charged residues comprising a potential casein kinase II phosphorylation site on serine residue 251 of TMX4. The combination of these two sequence features has been shown to constitute a potential binding site for PACS-2, a cytosolic sorting protein that uses COPI to target its substrate proteins to the ER dependent on their phosphorylation state (Kottgen et al., 2005; Simmen et al., 2005). While speculative at present, it is conceivable that the cellular localization and thereby the function of TMX4 and TMX is modulated by binding of regulatory sorting factors that target different binding sites in the cytosolic tail of the proteins. Preliminary *in vitro* experiments confirmed that PACS-2 interacts with TMX4 and more weakly with TMX (Thomas Simmen, personal communication). Concerning a potential functional connection, it is interesting to note that PACS-2 promotes staurosporin-induced apoptosis (Simmen 2005), while overexpression of TMX has been shown to protect cells from BFA-induced apoptosis (Matsuo et al., 2001). TMX4 knock-down has been found to enhance thapsigargin-induced apoptosis (Thomas Simmen, personal communication). We are currently pursuing the idea that the sorting motifs in TMX4 and TMX and a potential interaction with PACS-2 and other cytosolic factors connect the two proteins to a function in apoptosis.

4 Materials and Methods

Cell Culture, Transfection and Antibodies

Mammalian Expression Plasmids

Bacterial Expression Plasmids

Immunofluorescence Microscopy

Gel Electrophoresis and Western Blot

Pulse-Chase Experiments, Immunoprecipitation and Crosslinking

Endoglycosidase H Digestion

Membrane Association and Proteinase K Protection Assay

Tissue Northern Blot

Determination of the In Vivo Redox State

Protein Expression and Purification

Circular Dichroism Measurements

Cell Surface Biotinylation

Velocity Sedimentation on Sucrose Gradients

Cell Culture, Transfection and Antibodies

HeLa and Vero cells were maintained in 10 ml α-Minimum Essential Medium (MEM) (Invitrogen) supplemented with 10% fetal calf serum (FCS) at 37 °C under 5 % CO_2. Vero cells and the HeLa cells used in Fig. 2 were transfected using an AMAXA electroporator following the manufacturer's recommendations for these cell lines. For all other experiments, HeLa cells were transfected with Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The mouse monoclonal antibodies HA.11 (clone 16B12) against the HA epitope and clone 9E10 against the c-myc epitope were from Covance Research Products (Berkeley, CA, USA). The affinity purified rabbit antiserum TMX3-C specific for TMX3 has previously been described (Haugstetter et al., 2005). The polyclonal rabbit antiserum against calnexin was generously provided by A. Helenius, ETH Zurich. Alexa Fluor 594 goat anti-mouse and Alexa Fluor 488 goat anti-rabbit were from Molecular Probes (Leiden, The Netherlands). Horseradish peroxidase-coupled goat anti-mouse and rabbit IgG antibodies were from Pierce.

Mammalian Expression Plasmids

The TMX4 cDNA clone 11412-f01 (IMAGE clone ID: 5165952) was obtained from MRC Geneservice (Cambridge, UK). It was amplified using the primers C5-11412fo1-for* and C8-11412fo1-HArev* and inserted into the pcDNA3.1 vector (Stratagene) using the restriction sites NotI and XbaI to yield the construct pcDNA3.1/TMX4. Table 4 in the Appendix provides an overview of the sequences of all primers used in this study. All constructs were made by the cloning of PCR products into the pcDNA3.1 vector using the the restriction sites NotI and XbaI with the exception pcDNA3.1/TMX4-myc where we used XhoI and XbaI. The correct sequences of all generated plasmids were verified by DNA sequencing.

The construct encoding N-terminally HA-tagged TMX4, pcDNA3.1/HA-TMX4, was generated by overlap extension PCR. Here, the products from two initial separate PCR reactions (PCR(A) and PCR(B)) were used as templates in a third reaction (PCR(C)) together with the forward primer of reaction A and the reverse primer of reaction B. We used pcDNA3.1/TMX4 as a template for the two initial PCR reactions together with the primers C5-11412fo1-for* and C8-11412fo1-HArev* (PCR(A)) and C7-11412fo1-HAfor* and C6-11412fo1-rev (PCR(B)). The product of PCR(C) was inserted into pcDNA3.1 and the resulting construct, pcDNA3.1/HA-

TMX4, was used as a template to generate pcDNA3.1/HA-TMX-KQK by means of the QuikChange Site-Directed Mutagenesis Kit (Stratagene) using the primers "TMX4 R to K for" and "TMX4 R to K rev". As a result of the mutagenesis Arg338 and Arg340 are both exchanged for lysines.

The pcDNA3.1/HA-TMX4-CPSS construct contains a CPSS active-site sequence instead of the endogenous CPSC. pcDNA3.1/HA-TMX4 was used as a template for PCR(A) and (B) where the primer pairs C5-11412fo1-for* and "TMX4 C to S rev" (PCR(A)) and C6-11412fo1-rev and "TMX4 C to S for" (PCR(B)) were used. TMX4 devoid of the C-terminal tail, pcDNA3.1/Δtail, was generated using pcDNA3.1/HA-TMX4 as a template and the primers C5-11412fo1-for* and "TMX4 lum+TM rev".

To make the construct encoding C-terminally myc-tagged TMX4, pcDNA3.1/TMX4-myc, we first equipped pcDNA3.1/TMX4 with a XhoI restriction sites before the stop codon in a PCR reaction with the primers C5-11412fo1-for* and T4myc-Xhorev. The resulting construct was named pcDNA3.1/TMX4xhoI. Then, a triple myc tag from the pcDNA3/Ero1β–3myc construct (kindly provided by R. Sitia, Milan) was amplified using the primers 3xmycfor and 3xmycrev. The PCR product was cloned into pcDNA3.1/TMX4xhoI using XhoI and XbaI restriction sites.

Human CD4 cDNA (Maddon et al., 1985), kindly provided by H. P. Hauri, University of Basel (with the permission of K. Karjalainen), was subcloned into pcDNA3.1 after amplification using the primers C4T41for and C4T4F2rev to yield pcDNA3.1/CD4. Three chimeras between CD4 and TMX4 were then made by overlap extension PCR. PCR(A) was perfomed with pcDNA3.1/CD4 as template and C4T41for and C6-11412fo1-rev as primers. For PCR(B), pcDNA3.1/HA-TMX4 was used as template together with the primers C4T4F4for and C4T4F3rev. The product of PCR(C) was inserted into pcDNA3.1 to obtain pcDNA3.1/CD4-2. The construct pcDNA/CD4-3 was produced in a similar way, but using the primers C4T4F6for and C4T4F5rev in the PCR(B) reaction, and the primer pair C4T41for and C4T4F5rev for PCR(C). The plasmid pcDNA3.1/CD4-4 was generated as pcDNA/CD4-3, but with pcDNA3.1/HA-TMX4-KQK as a template in PCR(B).

The wild-type CD construct and the CD4-TMX4 chimeras were then HA tagged. PCR(A) was identical for all four, with pcDNA3.1/CD4 as template and the primers C4T4kozakfor and C4T4HArev. In PCR(B), C4T4HAfor was used as

forward primer and combined with different templates and reverse primers: i) for pcDNA3.1/HA-CD4, the template was pcDNA3.1/CD4 and the primer C4T4F2rev, ii) for pcDNA3.1/HA-$LTM_{CD4}C_{TMX4}$, the template was pcDNA3.1/CD4-2 and the primer C6-11412fo1-rev; iii) for pcDNA3.1/HA-$LTM_{CD4}C_{TMX4}$ the template was pcDNA3.1/CD4-3 and the primer C6-11412fo1-rev; iv) for pcDNA3.1/HA-$LTM_{CD4}C_{TMX4\text{-}KQK}$ the template was pcDNA3.1/CD4-4 and the primer C6-11412fo1-rev. For PCR(C), products A and B were combined with the forward primer C4T4kozakfor and the reverse primer corresponding to product B.
The primer sequences are listed in table 4 in the appendix.

Bacterial Expression Plasmids

All three constructs were cloned into the vector pRSET-A-MiniT by means of the restriction sites EcoRI and BamHI. For His-TMX4-Trx, the thiorexin-like domain was amplified in a PCR reaction using the TMX cDNA clone and the primers T3Tfor and T3Trev. The construct for His-TMX4-Lum was generated with the same template and forward primer but with T3Lrev as reverse primer. The sequence of the thioredoxin-like domain was cut out of the construct for His-TMX4-Trx by EcoRI and BamHI and inserted into the pRSET-A-miniT-His-GB1 vector, kindly provided by Kurt Wüthrich, ETH Zurich, to yield the construct for His-GB1-TMX4-Trx.

Immunofluorescence Microscopy

Transfected cells on microscope coverslips were fixed 24 h post-transfection using the method of McLean and Nakane (McLean and Nakane, 1974). For studying intracellular proteins, cells were incubated for 4 h in fixing solution (10 mM sodium meta-periodate, 75 mM lysine, 375 mM sodium phosphate, pH 6.2, 2% paraformaldehyde). The fixed cells were rinsed three times with PBS^{++} (phosphate-buffered saline; 154 mM NaCl, 1.9 mM KH_2PO_4, 8.1 mM Na_2HPO_4, pH 7.3, 1mM $CaCl_2$, 0.5 mM $MgCl_2$) at room temperature, blocked with a solution of 3% bovine serum albumine (BSA) in PBS^{++} for 15 min and treated with permeabilization solution (0.05% saponin, 3% BSA in PBS^{++}) for 20 min. Next, cells were incubated with the monoclonal HA.11 antibody, diluted 1:1000 in permeabilization buffer, or the polyclonal α-CNX, diluted 1:500 in permeabilization buffer. After incubation for 1 h at room temperatur, the cells were washed three times in permeabilization buffer, followed by incubation for 45 min with secondary antibody, diluted 1:200 in

permeabilization buffer. As secondary antibodies, we used Alexa 594 goat anti-mouse and Alexa 488 goat anti-rabbit. Coverslips were rinsed three times with permeabilization buffer, once with PBS^{++} and once briefly with water, and mounted on slides with Gel Mount™ Aqueous Mounting Medium (Sigma). After 1 h, the coverslip edges were resealed with Clarion™ Mounting Medium (Sigma). For immunoflourescence microscopy of cell surface exposed proteins on non-permeabilized cells, transfected cells were blocked and incubated with primary antibody on ice before fixation followed by incubation with secondary antibody and mounting. For colocalization with KDEL receptor, cells were treated with 10 µg/ml BFA for 1 h at 37 °C. Confocal images of fixed cells were acquired using either an inverted Zeiss LSM 510 metamicroscope (Carl Zeiss MicroImaging, Inc.) or a TCS NT/SP confocal laser scanning microscope (Leica Microsystems Heidelberg GmbH, Germany). Both microscopes were equipped with an apochrome 100X objective, NA 1.4.

Gel Electrophoresis and Western Blot

Samples were separated on 10% 10x10.5 cm Hoefer minigels (GE Healthcare), with the exception of the CD4-TMX4 chimeras where a concentration of 7.5% polyacrylamide was used. For analysis of sucrose gradient fractions, SDS-PAGE was performed in a Hoefer standard vertical unit (18x16 cm gels). Western blotting was performed as described in (Haugstetter et al., 2005) using the ECL Advance Western blotting detection system (GE Healthcare). Antibodies were used in the following dilutions; HA.11 (anti-HA) 1:1000; TMX3-C (anti-TMX3) 1:500; horseradish peroxidase-coupled goat secondary antibodies (anti-mouse and rabbit) 1:50,000. Band intensities obtained by autoradiography or Western blotting were quantified using ImageJ (Abramoff et al., 2004).

Pulse-Chase Experiments, Immunoprecipitation and Crosslinking

12 h after transfection, cells were washed once with starvation medium (Dulbecco's modified Eagle's medium without methionine and cystine from Sigma), incubated in starvation medium for 15 min at 37 °C, and pulsed for 5 min or overnight at 37 °C with 125 µCi [^{35}S]methionine and [^{35}S]cysteine (PerkinElmer) in 0.5 ml pulse medium per 35-mm dish. The cells were washed once with warm chase medium (α-MEM/10% FCS containing 5 mM methionine) and incubated in the same medium at

37 °C. At the end of the chase, dishes were transferred to ice where all subsequent steps were carried out. The cells were washed twice with PBS, scraped off with a rubber policeman in 1 ml lysis buffer (100 mM sodium phosphate, pH 8, 1% Triton X-100, and 1 mM phenylmethylsulfonyl fluoride (PMSF)), and passed 10 times through a 27 G needle. After incubation for 60 min, the lysate was centrifuged at 25,000 g for 30 min, and the supernatant was rotated for 2 h at 4 °C with 50 µl protein A-sepharose beads (GE Healthcare) preadsorbed with either 2 µl HA.11 or 2µl 9E10. After washing four times with lysis buffer and once with 100 mM sodium phosphate pH 8.0, proteins eluted from the beads by boiling in sample buffer. For crosslinking experiments, cells were incubated with either suberic acid bis(N-hydroxysuccinimide ester) (DSS, Sigma) or the sulfhydril-specific cross-linker bis-maleimidohexane (BMH, Pierce) and treated as described (Haugstetter et al., 2005).

Endoglycosidase H Digestion

HeLa lysates for Endo H digestions were either lysed in SDS sample buffer (0.2 M Tris-HCl, pH 6.8, 80 mg/ml SDS, 40% w/w glycerol, 80 ng/ml bromophenol blue, 10 mg/ml DTT) and denatured at 96 °C for 15 min (Fig. 2) or pulse/labeled, immunoprecipitated and released from the beads by denaturing at 96 °C for 15 min in loading buffer (Fig. 7). The samples were then supplemented with one-tenth volume of 50 mM sodium citrate, pH 5.5, and treated with endoglycosidase H (New England Biolabs) for 1 h at 37 °C.

Membrane association and Proteinase K Protection Assay

To study the membrane association of TMX4, HeLa cells at 60% confluency were harvested, fractionated and extracted in sodium carbonate as previously described (Haugstetter et al., 2005). To determine the topology of TMX4, HeLa cells were labeled with [^{35}S]methionine and [^{35}S]cysteine (PerkinElmer), washed twice with PBS (PBS^{++} without $CaCl_2$ and $MgCl_2$), scraped off in homogenization buffer (20 mM Hepes-NaOH pH 7.5, 0.25 M sucrose, 1 mM DTT) and passed 10 times through a 27 gauge needle. After removal of cell debris by centrifugation at 300 g for 3 min at 4 °C, the supernatant was centrifuged at 100,000 g for 1 h at 4 °C. The pellet was suspended in assay buffer (25 mM Tris-HCl, pH 8, 500 mM NaCl, 1 mM DTT) and treated with 60 µg/ml proteinase K (Roche Molecular Biochemicals) on ice for 1 h in the presence or absence of 1% Triton X-100. Proteinase K activity was inhibited with

1 mM PMSF and the samples were separated by SDS-PAGE and visualized by autoradiography.

Tissue Northern Blot

A human multiple tissue Northern blot (Ambion Inc, Austin, TX) was hybridized with radioactive probes generated from a NotI-XhoI fragment obtained from restriction digestion of pcDNA3.1/TMX4-myc. The gel-purified fragment was labeled with [α-^{32}P]dCTP using the Strip-EZ DNA Kit (Ambion), followed by purification using a Micro Bio-Spin 30 column (Bio-Rad Laboratories) to remove unincorporated nucleotides. Hybridization was done at 65 °C in ULTRAhyb buffer (Ambion) following the manufacturer's protocol. The probed blot was visualized by autoradiography. The blot was then stripped according to the Strip-EZ DNA Kit protocol and hybridized with probes derived from the β-actin mouse DNA template delivered with the kit.

Determination of the In Vivo Redox State

Metabolically [^{35}S]methionine/cysteine-labeled HeLa cells transfected with HA-TMX4 and HA-TMX4-CPSS were treated with dithiothreitol (DTT) and diamide to obtain reduced and oxidized control samples, modified in situ with N-ethylmaleimide (NEM), homogenized and lysed as described recently (Appenzeller-Herzog and Ellgaard, 2007). For the analysis of TMX3, cell lysates were used directly for alkylation by AMS (Appenzeller-Herzog and Ellgaard, 2007). For the TMX4 proteins, immunoprecipitated HA-TMX4 and HA-TMX4-CPSS were used as the starting material for AMS alkylation. Briefly, 27 µl of the sample were reduced with 1.5 µl 200 mM TCEP for 15 min and treated with 7.5 µl 75 mM AMS for 1h at room temperature. After separation by SDS-PAGE, TMX3 was visualized by Western blotting, and HA-TMX4 and HA-TMX4-CPSS by autoradiography.

Protein Expression and Purification

pRSET-A-miniT/TMX4-Trx was transformed into the BL21(DE3) pLysS *E. coli* strain and plated on LB-agar containing 100 µg/ml ampicillin. From a fresh transformation, one colony was picked to inoculate 100-200 ml LB medium

preculture that was grown at 37 °C for at least 12 h containing antibiotics as described above.

For test expressions at different temperatures, the expression of His-TMX4-Trx was induced with 1 mM isopropyl β-D-1-thiogalactopryranoside (IPTG), and the culture was grown for additional 3-4 h in an orbital shaker at 250 rpm, 37 °C before 500 µl cell suspension were taken and centrifuged for 5 min at 4,000 x g. The pellet was resuspended in 50 µl reducing SDS sample buffer, yielding the total lysate. In order to separate soluble from insoluble protein, 5 ml cell suspension were centrifuged for 5 min at 4,000 x g, the pellet was resuspended in 500 µl non-denaturing resuspension buffer (buffer A: 500 mM NaCl, 25 mM Tris-HCl, pH 8.0, 1 mM DTT) and the cells were ruptured by sonication. Insoluble protein was pelleted in 100 µl reducing SDS-sample buffer and 166 µl 4x SDS sample buffer were added to the supernatant. The fractions were analysed by SDS-PAGE.

For purification of His-TMX4-Trx, 4-6 l LB medium with antibiotics were inoculated with 1.5% of the preculture and grown at 250 rpm until a A_{600} between 0.5 and 0.6 was reached. The expression of His-TMX4-Trx was induced with 1 mM IPTG, and the culture was grown for additional 3-4 h at 37 °C before the cells were harvested by centrifugation in a Sorvall SLA-3000 rotor at 4,000 rpm. Cell pellets were resuspended in denaturing resuspension buffer (buffer B: 6 M guanidine hydrochloride, 500 mM NaCl, 25 mM Tris-HCl, pH 8.0, 1 mM DTT), sonicated and subjected to centrifugation for 45 min using an SS-34 rotor (Sorvall) at 18,000 rpm. The clear supernatant was applied to a Ni^{2+}-charged nitrilotriacetic acid metal affinity chromatography column (Qiagen) that was equilibrated with 5 column volumes of buffer B. After loading and washing with at least five column volumes of buffer B, the protein was refolded while bound on the column using a linear gradient over 100 ml from buffer B into buffer A. Subsequently, the refolded protein was eluted with a linear gradient over 10 column volumes into non-denaturing imidazole elution buffer (buffer C: corresponds to buffer B containing 500 mM imidazol). To elute protein that aggregated on the column, 100% EDTA-urea buffer (buffer D: 8 M urea, 25 mM Tris-HCl, pH 8.0, 25 mM EDTA) were applied until A_{280} was stable. Fractions containing His-TMX4-Trx were pooled and refolded by dialysis or rapid dilution.

For refolding by rapid dilution the protein was transferred into a 10 x excess of buffer A under vigorous stirring at 4 °C. After several hours of stirring at 4 °C, the protein solution was concentrated and checked for precipitation.

After dialysis into low salt buffer (buffer E: 25 mM NaCl, 25 mM Tris-HCl, pH 8.0, 1 mM DTT), the proteins were loaded onto a SOURCE15Q anion exchange chromatography column (Amersham Biosciences) eqilibrated in buffer E and eluted with a linear gradient over 100 ml against high salt buffer (buffer F: 500 mM NaCl, 25 mM Tris-HCl, pH 8.0, 1 mM DTT). Protein that aggregated on the column was eluted with 100 % denaturing high salt buffer (buffer G: corresponds buffer F with 8 M urea) until A_{280} had reached a constant level.

Expression and purification of His-GB1-TMX4-Trx was carried out as for His-TMX4-Trx. To remove the GB1 domain, the pooled fractions containing the pure protein were dialyzed into buffer A without DTT. 1 ml of the dialyzed protein solution at 0.16 mg/ml was incubated with 0.5 µg/ml Factor X (1 U/µl, Merck) at room temprature. Samples were taken after 3.5 h and 20 h and analyzed by SDS-PAGE. The reaction mixture was checked for precipitation by eye.

Circular Dichroism Measurements

CD measurements were performed on a Jasco J-810 spectropolarimeter in a 1-mm path length cell at 4 °C and at 95 °C. Spectra were measured between wavelengths of 200 and 260 nm, averaged from 3 scan, and the buffer base line was subtracted. His-TMX4-Trx was used at a concentration of 6.8 µM in a buffer containing 100 mM KH_2PO_4/KOH, pH 7.0 and 1 mM DTT. For temperature transition experiments, His-TMX4-Trx was measured at 220 nm in a temperature gradient from 4 °C and 85 °C increasing 1 °C per minute. A sample point was taken each 0.2 °C. After reaching 85 °C, the sample was allowed to cool down at 4 °C for 1 h and the measurement was repeated.

Cell Surface Biotinylation

24 h after transfection, cells were rinsed three times with ice-cold PBS^{++} and incubated with 0.5 mg/ml sulfo-NHS-LC-LC-biotin (Pierce) in PBS^{++} for 30 min on ice. The biotinylation reaction was stopped by washing once with PBS^{++} and twice with PBS^{++} containing 15 mM glycine, before the cells were scraped off and lysed in 500 µl NP-40 lysis buffer (10 mM Tris-HCl, pH 7.4, 1% Nonidet P-40, 0.4% sodium deoxycholate, 66 mM EDTA, 1 mM phenlymethylsulfonyl fluoride (PMSF)) by 10 passages through a 25 G needle and 1 h incubation on ice. After clearing by centrifugation for 10 min at 20,800 x g and 4 °C, 5 % of the lysate were boiled with

SDS sample buffer and kept as "total lysate" sample. To isolate cell surface proteins, the cleared lysate was supplemented with SDS to a final concentration of 0.3% and incubated with Strep-Tactin beads (Sigma) for 2 h at 4 °C on an end-over-end rotator after 50 µl of a 50% slurry of Strep-Tactin beads per sample had been preincubated in NP-40 lysis buffer containing 1:1000 CLAP proteasome inhibitor mix (Sigma). The beads were washed six times with NP-40 lysis buffer, once with NP-40 lysis buffer containing 0.3% SDS and 0.5 M NaCl and once with PBS^{++}, before boiling in 4x SDS sample buffer to elute the proteins. Total lysate and cell surface samples were analyzed by SDS-PAGE and Western blot.

Velocity Sedimentation on Sucrose Gradients

HeLa cells were lysed with 2% CHAPS and 1% Triton X-100 in HBS buffer (50 mM HEPES pH 7.5, 0.2 M NaCl). The lysates were then resolved on 5 ml 5-25% continuous sucrose gradiens in HBS supplemented with 0.1% Triton X-100, and centrifuged in a Sorvall AH-650 rotor at 42,000 rpm for 16 h at 4°C. The two standard proteins, BSA and β-amylase, were resolved in the same fashion. After the centrifugation, the gradients were fractionated manually from top to bottom. Each fraction containing HA-TMX4 was analyzed by Western blotting, followed by quantification of band intensities as described above. The protein content in fractions containing standard proteins was determined using the Coomassie Plus Bradford Assay Kit (Pierce).

5 References

Abramoff, M.D., Magelhaes, P.J. and Ram, S.J. (2004) Image Processing with ImageJ. *Biophotonics International*, **11**, 36-42.

Appenzeller-Herzog, C. and Ellgaard, L. (2007) Determination of the in vivo reduction-oxidation state of protein disulfide isomerase. *Submitted.*

Baryshev, M., Sargsyan, E. and Mkrtchian, S. (2006) ERp29 is an essential endoplasmic reticulum factor regulating secretion of thyroglobulin. *Biochem Biophys Res Commun*, **340**, 617-624.

Bass, R., Ruddock, L.W., Klappa, P. and Freedman, R.B. (2004) A major fraction of endoplasmic reticulum-located glutathione is present as mixed disulfides with protein. *J Biol Chem*, **279**, 5257-5262.

Bessette, P.H., Qiu, J., Bardwell, J.C., Swartz, J.R. and Georgiou, G. (2001) Effect of sequences of the active-site dipeptides of DsbA and DsbC on in vivo folding of multidisulfide proteins in Escherichia coli. *J Bacteriol*, **183**, 980-988.

Card, P.B. and Gardner, K.H. (2005) *Identification and Optimization of Protein Domains for NMR Studies.*

Cartier, E.A., Conti, L.R., Vandenberg, C.A. and Shyng, S.L. (2001) Defective trafficking and function of KATP channels caused by a sulfonylurea receptor 1 mutation associated with persistent hyperinsulinemic hypoglycemia of infancy. *Proc Natl Acad Sci U S A*, **98**, 2882-2887.

Chang, X.B., Cui, L., Hou, Y.X., Jensen, T.J., Aleksandrov, A.A., Mengos, A. and Riordan, J.R. (1999) Removal of multiple arginine-framed trafficking signals overcomes misprocessing of delta F508 CFTR present in most patients with cystic fibrosis. *Mol Cell*, **4**, 137-142.

Chivian, D., Kim, D.E., Malmstrom, L., Schonbrun, J., Rohl, C.A. and Baker, D. (2005) Prediction of CASP6 structures using automated Robetta protocols. *Proteins*, **61 Suppl 7**, 157-166.

Cosson, P. and Letourneur, F. (1994) Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. *Science*, **263**, 1629-1631.

Danilczyk, U.G., Cohen-Doyle, M.F. and Williams, D.B. (2000) Functional relationship between calreticulin, calnexin, and the endoplasmic reticulum luminal domain of calnexin. *Journal of Biological Chemistry*, **275**, 13089-13097.

Darby, N.J. and Creighton, T.E. (1995) Characterization of the active site cysteine residues of the thioredoxin-like domains of protein disulfide isomerase. *Biochemistry*, **34**, 16770-16780.

Ellgaard, L. and Helenius, A. (2003) Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol*, **4**, 181-191.

Ellgaard, L. and Ruddock, L.W. (2005) The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep.*, **6**, 28-32.

Elliott, J.G., Oliver, J.D. and High, S. (1997) The thiol-dependent reductase ERp57 interacts specifically with N- glycosylated integral membrane proteins. *J. Biol. Chem.*, **272**, 13849-13855.

Ensembl. release 46 - August 2007. www.ensembl.org.

Eugster, A., Frigerio, G., Dale, M. and Duden, R. (2004) The alpha- and beta'-COP WD40 domains mediate cargo-selective interactions with distinct di-lysine motifs. *Mol Biol Cell*, **15**, 1011-1023.

Felsenstein, J. (1989) Phylogeny Inference Package (Version 3.2). *Cladistics*, **5**, 164-166.

Forman-Kay, J.D., Clore, G.M., Wingfield, P.T. and Gronenborn, A.M. (1991) High-resolution three-dimensional structure of reduced recombinant human thioredoxin in solution. *Biochemistry*, **30**, 2685-2698.

Frickel, E.M., Frei, P., Bouvier, M., Stafford, W.F., Helenius, A., Glockshuber, R. and Ellgaard, L. (2004) ERp57 Is a Multifunctional Thiol-Disulfide Oxidoreductase. *J. Biol. Chem.*, **279**, 18277-18287.

Gaynor, E.C., te Heesen, S., Graham, T.R., Aebi, M. and Emr, S.D. (1994) Signal-mediated retrieval of a membrane protein from the Golgi to the ER in yeast. *J Cell Biol*, **127**, 653-665.

Girod, A., Storrie, B., Simpson, J.C., Johannes, L., Goud, B., Roberts, L.M., Lord, J.M., Nilsson, T. and Pepperkok, R. (1999) Evidence for a COP-I-independent transport route from the Golgi complex to the endoplasmic reticulum. *Nat Cell Biol*, **1**, 423-430.

Goldberger, R.F., Epstein, C.J. and Anfinsen, C.B. (1964) Purification and Properties of a Microsomal Enzyme System Catalyzing the Reactivation of Reduced Ribonuclease and Lysozyme. *J Biol Chem*, **239**, 1406-1410.

Grauschopf, U., Winther, J.R., Korber, P., Zander, T., Dallinger, P. and Bardwell, J.C. (1995) Why is DsbA such an oxidizing disulfide catalyst? *Cell*, **83**, 947-955.

Griffiths, G., Ericsson, M., Krijnse-Locker, J., Nilsson, T., Goud, B., Soling, H.D., Tang, B.L., Wong, S.H. and Hong, W. (1994) Localization of the Lys, Asp, Glu, Leu tetrapeptide receptor to the Golgi complex and the intermediate compartment in mammalian cells. *J Cell Biol*, **127**, 1557-1574.

Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*, **52**, 696-704.

Hardt, B., Kalz-Fuller, B., Aparicio, R., Volker, C. and Bause, E. (2003) (Arg)3 within the N-terminal domain of glucosidase I contains ER targeting information but is not required absolutely for ER localization. *Glycobiology*, **13**, 159-168.

Haugstetter, J., Blicher, T. and Ellgaard, L. (2005) Identification and Characterization of a Novel Thioredoxin-related Transmembrane Protein of the Endoplasmic Reticulum. *J. Biol. Chem.*, **280**, 8371-8380.

Haugstetter, J., Maurer, M.A., Blicher, T., Pagac, M., Wider, G. and Ellgaard, L. (2007) Structure-function analysis of the endoplasmic reticulum oxidoreductase TMX3 reveals interdomain stabilization of the N-terminal redox-active domain. *J. Biol. Chem.*, **282**, 33859-33867.

Hermosilla, R., Oueslati, M., Donalies, U., Schonenberger, E., Krause, E., Oksche, A., Rosenthal, W. and Schulein, R. (2004) Disease-causing V(2) vasopressin receptors are retained in different compartments of the early secretory pathway. *Traffic*, **5**, 993-1005.

Higo, T., Hattori, M., Nakamura, T., Natsume, T., Michikawa, T. and Mikoshiba, K. (2005) Subtype-specific and ER lumenal environment-dependent regulation of inositol 1,4,5-trisphosphate receptor type 1 by ERp44. *Cell*, **120**, 85-98.

Hol, W.G. (1985) The role of the alpha-helix dipole in protein function and structure. *Prog Biophys Mol Biol*, **45**, 149-195.

Huber-Wunderlich, M. and Glockshuber, R. (1998) A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Fold Des*, **3**, 161-171.

Huth, J.R., Bewley, C.A., Jackson, B.M., Hinnebusch, A.G., Clore, G.M. and Gronenborn, A.M. (1997) Design of an expression system for detecting folded

protein domains and mapping macromolecular interactions by NMR. *Protein Sci*, **6**, 2359-2364.

Hwang, C., Sinskey, A.J. and Lodish, H.F. (1992) Oxidized redox state of glutathione in the endoplasmic reticulum. *Science*, **257**, 1496-1502.

Itin, C., Schindler, R. and Hauri, H.P. (1995) Targeting of protein ERGIC-53 to the ER/ERGIC/cis-Golgi recycling pathway. *J Cell Biol*, **131**, 57-67.

Jackson, M.R., Nilsson, T. and Peterson, P.A. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *Embo J*, **9**, 3153-3162.

Jackson, M.R., Nilsson, T. and Peterson, P.A. (1993) Retrieval of transmembrane proteins to the endoplasmic reticulum. *J Cell Biol*, **121**, 317-333.

Jeng, M.F., Campbell, A.P., Begley, T., Holmgren, A., Case, D.A., Wright, P.E. and Dyson, H.J. (1994) High-resolution solution structures of oxidized and reduced Escherichia coli thioredoxin. *Structure*, **2**, 853-868.

Johansson, E., Majka, J. and Burgers, P.M. (2001) Structure of DNA polymerase delta from Saccharomyces cerevisiae. *J Biol Chem*, **276**, 43824-43828.

Klappa, P., Freedman, R.B. and Zimmermann, R. (1995) Protein disulphide isomerase and a lumenal cyclophilin-type peptidyl prolyl cis-trans isomerase are in transient contact with secretory proteins during late stages of translocation. *Eur J Biochem*, **232**, 755-764.

Ko, F.C. and Chow, K.L. (2002) A novel thioredoxin-like protein encoded by the C. elegans dpy-11 gene is required for body and sensory organ morphogenesis. *Development*, **129**, 1185-1194.

Kortemme, T. and Creighton, T.E. (1995) Ionisation of cysteine residues at the termini of model alpha-helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *J Mol Biol*, **253**, 799-812.

Kottgen, M., Benzing, T., Simmen, T., Tauber, R., Buchholz, B., Feliciangeli, S., Huber, T.B., Schermer, B., Kramer-Zucker, A., Hopker, K., Simmen, K.C., Tschucke, C.C., Sandford, R., Kim, E., Thomas, G. and Walz, G. (2005) Trafficking of TRPP2 by PACS proteins represents a novel mechanism of ion channel regulation. *Embo J*, **24**, 705-716.

Krause, G., Lundstrom, J., Barea, J.L., Pueyo de la Cuesta, C. and Holmgren, A. (1991) Mimicking the active site of protein disulfide-isomerase by substitution of proline 34 in Escherichia coli thioredoxin. *J. Biol. Chem.*, **266**, 9494-9500.

Lappi, A.K., Lensink, M.F., Alanen, H.I., Salo, K.E., Lobell, M., Juffer, A.H. and Ruddock, L.W. (2004) A conserved arginine plays a role in the catalytic cycle of the protein disulphide isomerases. *J. Mol. Biol.*, **335**, 283-295.

Lee, S.K. (1999) Four consecutive arginine residues at positions 836-839 of EBV gp110 determine intracellular localization of gp110. *Virology*, **264**, 350-358.

Letourneur, F., Gaynor, E.C., Hennecke, S., Demolliere, C., Duden, R., Emr, S.D., Riezman, H. and Cosson, P. (1994) Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell*, **79**, 1199-1207.

Letourneur, F., Hennecke, S., Demolliere, C. and Cosson, P. (1995) Steric masking of a dilysine endoplasmic reticulum retention motif during assembly of the human high affinity receptor for immunoglobulin E. *J Cell Biol*, **129**, 971-978.

Lewis, M.J. and Pelham, H.R. (1990) A human homologue of the yeast HDEL receptor. *Nature*, **348**, 162-163.

Lewis, M.J. and Pelham, H.R. (1992a) Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum. *Cell*, **68**, 353-364.

Lewis, M.J. and Pelham, H.R. (1992b) Sequence of a second human KDEL receptor. *J Mol Biol*, **226**, 913-916.

Lippincott-Schwartz, J., Donaldson, J.G., Schweizer, A., Berger, E.G., Hauri, H.P., Yuan, L.C. and Klausner, R.D. (1990) Microtubule-dependent retrograde transport of proteins into the ER in the presence of brefeldin A suggests an ER recycling pathway. *Cell*, **60**, 821-836.

Ma, D. and Jan, L.Y. (2002) ER transport signals and trafficking of potassium channels and receptors. *Curr Opin Neurobiol*, **12**, 287-292.

Ma, D., Zerangue, N., Lin, Y.F., Collins, A., Yu, M., Jan, Y.N. and Jan, L.Y. (2001) Role of ER export signals in controlling surface potassium channel numbers. *Science*, **291**, 316-319.

Ma, D., Zerangue, N., Raab-Graham, K., Fried, S.R., Jan, Y.N. and Jan, L.Y. (2002) Diverse trafficking patterns due to multiple traffic motifs in G protein-activated inwardly rectifying potassium channels from brain and heart. *Neuron*, **33**, 715-729.

Maddon, P.J., Littman, D.R., Godfrey, M., Maddon, D.E., Chess, L. and Axel, R. (1985) The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family. *Cell*, **42**, 93-104.

Magnuson, B., Rainey, E.K., Benjamin, T., Baryshev, M., Mkrtchian, S. and Tsai, B. (2005) ERp29 triggers a conformational change in polyomavirus to stimulate membrane binding. *Mol Cell*, **20**, 289-300.

Margeta-Mitrovic, M., Jan, Y.N. and Jan, L.Y. (2000) A trafficking checkpoint controls GABA(B) receptor heterodimerization. *Neuron*, **27**, 97-106.

Martin, J.L. (1995) Thioredoxin--a fold for all reasons. *Structure*, **3**, 245-250.

Matsuo, Y., Akiyama, N., Nakamura, H., Yodoi, J., Noda, M. and Kizaka-Kondoh, S. (2001) Identification of a novel thioredoxin-related transmembrane protein. *J. Biol. Chem.*, **276**, 10032-10038.

Matsuo, Y., Nishinaka, Y., Suzuki, S., Kojima, M., Kizaka-Kondoh, S., Kondo, N., Son, A., Sakakura-Nishiyama, J., Yamaguchi, Y., Masutani, H., Ishii, Y. and Yodoi, J. (2004) TMX, a human transmembrane oxidoreductase of the thioredoxin family: the possible role in disulfide-linked protein folding in the endoplasmic reticulum. *Arch. Biochem. Biophys.*, **423**, 81-87.

McLean, I.W. and Nakane, P.K. (1974) Periodate-lysine-paraformaldehyde fixative. A new fixation for immunoelectron microscopy. *J Histochem Cytochem*, **22**, 1077-1083.

Meng, X., Zhang, C., Chen, J., Peng, S., Cao, Y., Ying, K., Xie, Y. and Mao, Y. (2003) Cloning and identification of a novel cDNA coding thioredoxin-related transmembrane protein 2. *Biochem. Genet.*, **41**, 99-106.

Meyer, G., Gicklhorn, D., Strive, T., Radsak, K. and Eickmann, M. (2002) A three-residue signal confers localization of a reporter protein in the inner nuclear membrane. *Biochem Biophys Res Commun*, **291**, 966-971.

Meyer, G.A. and Radsak, K.D. (2000) Identification of a novel signal sequence that targets transmembrane proteins to the nuclear envelope inner membrane. *Journal of Biological Chemistry*, **275**, 3857-3866.

Michelsen, K., Yuan, H. and Schwappach, B. (2005) Hide and run. Arginine-based endoplasmic-reticulum-sorting motifs in the assembly of heteromultimeric membrane proteins. *EMBO Rep*, **6**, 717-722.

Munro, S. and Pelham, H.R. (1987) A C-terminal signal prevents secretion of luminal ER proteins. *Cell*, **48**, 899-907.

Muslin, A.J. and Xing, H. (2000) 14-3-3 proteins: regulation of subcellular localization by molecular interference. *Cell Signal*, **12**, 703-709.

Nilsson, T., Jackson, M. and Peterson, P.A. (1989) Short cytoplasmic sequences serve as retention signals for transmembrane proteins in the endoplasmic reticulum. *Cell*, **58**, 707-718.

Nufer, O. and Hauri, H.P. (2003) ER export: call 14-3-3. *Curr. Biol.*, **13**, R391-393.

Nufer, O., Mitrovic, S. and Hauri, H.P. (2003) Profile-based data base scanning for animal L-type lectins and characterization of VIPL, a novel VIP36-like endoplasmic reticulum protein. *J. Biol. Chem.*, **278**, 15886-15896.

O'Kelly, I., Butler, M.H., Zilberberg, N. and Goldstein, S.A. (2002) Forward transport. 14-3-3 binding overcomes retention in endoplasmic reticulum by dibasic signals. *Cell*, **111**, 577-588.

Oliver, J.D., van der Wal, F.J., Bulleid, N.J. and High, S. (1997) Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science*, **275**, 86-88.

Orci, L., Perrelet, A., Ravazzola, M., Amherdt, M., Rothman, J.E. and Schekman, R. (1994) Coatomer-rich endoplasmic reticulum. *Proc Natl Acad Sci U S A*, **91**, 11924-11928.

Pelham, H.R. (1988) Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. *Embo J*, **7**, 913-918.

Pelham, H.R. (1989) Control of protein exit from the endoplasmic reticulum. *Annu Rev Cell Biol*, **5**, 1-23.

Pelham, H.R. (1990) The retention signal for soluble proteins of the endoplasmic reticulum. *Annu Rev Cell Biol*, **5**, 1-23.

Philipps, B. and Glockshuber, R. (2002) Randomization of the entire active-site helix alpha 1 of the thiol-disulfide oxidoreductase DsbA from Escherichia coli. *J Biol Chem*, **277**, 43050-43057.

Presley, J.F., Smith, C., Hirschberg, K., Miller, C., Cole, N.B., Zaal, K.J. and Lippincott-Schwartz, J. (1998) Golgi membrane dynamics. *Mol Biol Cell*, **9**, 1617-1626.

Quan, S., Schneider, I., Pan, J., Hacht, A.V. and Bardwell, J.C. (2007) The CXXC motif is more than a redox rheostat. *J Biol Chem.*

Raykhel, I., Alanen, H.I., Salo, K.E., Jurvansuu, J., Nguyen, V.D., Latva-Rana, M. and Ruddock, L.W. (2007) A molecular specificity code for the three mammalian KDEL receptors. *J. Cell Biol.*, **179**, 1193-1204.

Roche, K.W., Tu, J.C., Petralia, R.S., Xiao, B., Wenthold, R.J. and Worley, P.F. (1999) Homer 1b regulates the trafficking of group I metabotropic glutamate receptors. *J Biol Chem*, **274**, 25953-25957.

Roizman, B. and Furlong, D. (1974) *The replication of herpesvirus*. Plenum Press, Inc., New York, NY.

Rolls, M.M., Stein, P.A., Taylor, S.S., Ha, E., McKeon, F. and Rapoport, T.A. (1999) A visual screen of a GFP-fusion library identifies a new type of nuclear envelope membrane protein. *Journal of Cell Biology*, **146**, 29-43.

Sali, D., Bycroft, M. and Fersht, A.R. (1988) Stabilization of protein structure by interaction of alpha-helix dipole with a charged side chain. *Nature*, **335**, 740-743.

Sargsyan, E., Baryshev, M., Szekely, L., Sharipo, A. and Mkrtchian, S. (2002) Identification of ERp29, an endoplasmic reticulum lumenal protein, as a new member of the thyroglobulin folding complex. *J Biol Chem*, **277**, 17009-17015.

Schroder-Kohne, S., Letourneur, F. and Riezman, H. (1998) Alpha-COP can discriminate between distinct, functional di-lysine signals in vitro and regulates access into retrograde transport. *J Cell Sci*, **111 (Pt 23)**, 3459-3470.

Schutze, M.P., Peterson, P.A. and Jackson, M.R. (1994) An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum. *Embo J*, **13**, 1696-1705.

Schweizer, A., Fransen, J.A.M., Bächi, T., Ginsel, L. and Hauri, H.P. (1988) Identification,by a monoclonal antibody, of a 53-kD protein associates with a tubulo-vesicular compartment at the cis-side of the Golgi apparatus. *Journal of Cell Biology*, **107**, 1643-1653.

Schweizer, A., Fransen, J.A.M., Matter, K., Kreis, T.E., Ginsel, L. and Hauri, H.-P. (1990) The isolated ER-Golgi intermediate compartment involved in protein transport from the ER to the Golgi apparatus. *Eur J Cell Biol*, **53**, 185-196.

Scott, D.B., Blanpied, T.A. and Ehlers, M.D. (2003) Coordinated PKA and PKC phosphorylation suppresses RXR-mediated ER retention and regulates the surface delivery of NMDA receptors. *Neuropharmacology*, **45**, 755-767.

Scott, D.B., Blanpied, T.A., Swanson, G.T., Zhang, C. and Ehlers, M.D. (2001) An NMDA receptor ER retention signal regulated by phosphorylation and alternative splicing. *J Neurosci*, **21**, 3063-3072.

Serrano, L. and Fersht, A.R. (1989) Capping and alpha-helix stability. *Nature*, **342**, 296-299.

Shikano, S. and Li, M. (2003) Membrane receptor trafficking: evidence of proximal and distal zones conferred by two independent endoplasmic reticulum localization signals. *Proc Natl Acad Sci U S A*, **100**, 5783-5788.

Simmen, T., Aslan, J.E., Blagoveshchenskaya, A.D., Thomas, L., Wan, L., Xiang, Y., Feliciangeli, S.F., Hung, C.H., Crump, C.M. and Thomas, G. (2005) PACS-2 controls endoplasmic reticulum-mitochondria communication and Bid-mediated apoptosis. *Embo J*, **24**, 717-729.

Sonnhammer, E.L., von Heijne, G. and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol*, **6**, 175-182.

Standley, S., Roche, K.W., McCallum, J., Sans, N. and Wenthold, R.J. (2000) PDZ domain suppression of an ER retention signal in NMDA receptor NR1 splice variants. *Neuron*, **28**, 887-898.

Stockklausner, C., Ludwig, J., Ruppersberg, J.P. and Klocker, N. (2001) A sequence motif responsible for ER export and surface expression of Kir2.0 inward rectifier K(+) channels. *FEBS Lett*, **493**, 129-133.

Tang, B.L., Low, S.H., Hauri, H.P. and Hong, W. (1995) Segregation of ERGIC53 and the mammalian KDEL receptor upon exit from the 15 degrees C compartment. *Eur J Cell Biol*, **68**, 398-410.

Teasdale, R.D. and Jackson, M.R. (1996) Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the golgi apparatus. *Annu Rev Cell Dev Biol*, **12**, 27-54.

Tochio, N., Koshiba, S., Inoue, M., Kigawa, T. and Yokoyama, S. (2005) The solution structure of the thioredoxin-like domain of human Thioredoxin-related transmembrane protein. *To be Published*

Trowbridge, I.S., Collawn, J.F. and Hopkins, C.R. (1993) Signal-dependent membrane protein trafficking in the endocytic pathway. *Annu Rev Cell Biol*, **9**, 129-161.

Tzivion, G. and Avruch, J. (2002) 14-3-3 proteins: active cofactors in cellular regulation by serine/threonine phosphorylation. *J Biol Chem*, **277**, 3061-3064.

van Hemert, M.J., Steensma, H.Y. and van Heusden, G.P. (2001) 14-3-3 proteins: key regulators of cell division, signalling and apoptosis. *Bioessays*, **23**, 936-946.

Walter, S., Hubner, B., Hahn, U. and Schmid, F.X. (1995) Destabilization of a protein helix by electrostatic interactions. *J Mol Biol*, **252**, 133-143.

Wang, X., Matteson, J., An, Y., Moyer, B., Yoo, J.S., Bannykh, S., Wilson, I.A., Riordan, J.R. and Balch, W.E. (2004) COPII-dependent export of cystic fibrosis transmembrane conductance regulator from the ER uses a di-acidic exit code. *J Cell Biol*, **167**, 65-74.

Wieland, F. and Harter, C. (1999) Mechanisms of vesicle formation: insights from the COP system. *Curr Opin Cell Biol*, **11**, 440-446.

Xia, H., Hornby, Z.D. and Malenka, R.C. (2001) An ER retention signal explains differences in surface expression of NMDA and AMPA receptor subunits. *Neuropharmacology*, **41**, 714-723.

Yuan, H., Michelsen, K. and Schwappach, B. (2003) 14-3-3 dimers probe the assembly status of multimeric membrane proteins. *Curr Biol*, **13**, 638-646.

Zerangue, N., Jan, Y.N. and Jan, L.Y. (2000) An artificial tetramerization domain restores efficient assembly of functional Shaker channels lacking T1. *Proc Natl Acad Sci U S A*, **97**, 3591-3595.

Zerangue, N., Malan, M.J., Fried, S.R., Dazin, P.F., Jan, Y.N., Jan, L.Y. and Schwappach, B. (2001) Analysis of endoplasmic reticulum trafficking signals by combinatorial screening in mammalian cells. *Proc. Natl. Acad. Sci. USA*, **98**, 2431-2436.

Zerangue, N., Schwappach, B., Jan, Y.N. and Jan, L.Y. (1999) A new ER trafficking signal regulates the subunit stoichiometry of plasma membrane K(ATP) channels. *Neuron*, **22**, 537-548.

Zhou, P., Lugovskoy, A.A. and Wagner, G. (2001) A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. *J Biomol NMR*, **20**, 11-14.

Curriculum Vitae

DORIS ROTH

Thorsgade 62, 3.
DK-2200 Copenhagen N
Phone: +45-3011 0754
Email: doris.roth@bc.biol.ethz.ch

EDUCATION

PhD	**ETH Zurich**, Institute of Biochemistry, research advisors Prof. Lars Ellgaard and Prof. Ari Helenius.	expected 10/2007
Diploma	**ETH Zurich**, Biology with specialization in Biochemistry and Cell Biology.	10/2003
Preliminary Diploma	**Friedrich-Schiller-University Jena**, Germany, Biochemistry.	10/2000

FELLOWSHIPS

ETH Research Grant	06/2004-06/2007
Novartis Research Grant	01/2004-05/2004
Erasmus Exchange Fellowship	09/2000-09/2001

SCIENTIFIC EXPERIENCE

Doctoral work	**ETH Zurich/University of Copenhagen,** Denmark, research advisors Prof. Lars Ellgaard (Copenhagen) and Prof. Ari Helenius (Zurich): *TMX4 – a human thioredoxin-like protein in the ER.*	11/2003-present
Diploma Thesis	**Yale University, New Haven, USA**, research advisor Prof. Norma W. Andrews (New Haven) and Prof. Ari Helenius (ETH Zurich): *The Resealing Capacity of Primary Skin Fibroblasts with Genetic Defects Affecting Lysosomes.*	12/2002-06/2003
Semester work	**ETH Zurich**, research advisor Prof. Ari Helenius: *Cotranslational Folding of a Capsid Protease in Living Cells.*	10/2001-01/2002
Project internship	**ETH Zurich**, research advisor Prof. Nenad Ban: *Preparation and Crystallization of Eukaryotic Ribosomes.*	07 - 08/2001
Project internship	**University of Zurich**, advisor Prof. Andreas Plückthun: *Ribosome Display: Molecular Evolution and Selection.*	02 - 03/2001
Practical training	**Institute of Molecular Biotechnology Jena**, research advisor Prof. Frank Grosse.	08 - 09/2000

PUBLICATIONS

Roth D, Lynes E, Althaus N, Simmen T, Ellgaard L, manuscript in preparation: *A non-classical Localization Signal Targets the Novel ER Oxidoreductase TMX4 to the Endoplasmic Reticulum.*

Huynh C, Roth D, Ward DM, Kaplan J, Andrews NW (2004) PNAS 101(48): 16795-800: *Defective lysosomal exocytosis and plasma membrane repair in Chediak-Higashi/beige cells.*

Acknowledgements

Tusind tak to Lars for being a very trustworthy person in more than one way. Trustworthy in science because you are very precise and careful with results and their interpretation. That was what I was looking for in a supervisor and one of the reasons I wanted to work with you. Trustworthy outside science because I could always turn to you if there were problems of any kind. I appreciate a lot that you are (almost) always kind and patient and find it important to be a good group leader. Many thanks for changing my life by getting me to Denmark and making me like beer.

Tack så mycket to Ari for idealism and enthusiasm in science, the notion of a big scientific family with interesting good friends all over the world and science as a nationality. I was looking forward to discussions with you because I had the impression that you allow people to be creative and follow up ideas. That is one of the activities that are most exciting and fun to me. Thank you SO much for sending me to Yale – it was one of the best times of my life and a lasting experience.

Merci beaucoup and Merci vielmal to Yves Barral and Hans-Peter Hauri for being my thesis committee members – that involves a lot of work and in part traveling. I appreciate your commitment very much and felt both happy and lucky to have the opportunity to count on your ideas and advice in my PhD project.

Many thanks to LE group for a lot of help, information, ideas and protocols. I am thankful to my students Nils and Kathrin who not only established protocols and provided data but taught me at least as much about communication and teaching as I taught them about working in the lab. It was great to have someone to work on TMX4 together with me.

Vielen Dank to the Helenius group for being an unfailing source for material, information, help, friends, and companions in fortune and misfortune. I really missed you in Denmark. Thanks to the Protein Biology Group in Copenhagen for being sweet and helpful. I am thankful to Rudolf Glockshuber for providing facilities for CD measurements. The IBM in Zurich has an amazing staff. I am deeply indepted to Marianne, Toni, Katrin, Sonja and colleagues.

I greatly appreciate being a member of the MLS program because it created a network of fellow students (and friends) with an almost American atmosphere. As an extravert person I totally depended on my family and friends and Martin. I am very grateful for your support. I am very lucky because you are the greatest!

Appendix

Name	SwissProt accession	Length	ER retention	N-glycosylation sites (putative)	a-like domains
PDI	P07237	508	KDEL	0	2
ERp57	P30101	505	QDEL	0	2
PDIp	Q13087	525	KEEL	3	2
ERp72	P13667	645	KEEL	0	3
ERp65	Q8N807	584	KEEL	8	2
ERp27	Q96DN0	273	KVEL	1	0
PDIr	Q14554	519	KEEL	0	3
ERp28	P30040	261	KEEL	1	0
ERdj5	Q8IXB1	793	KDEL	1	4
P5	Q15084	440	KDEL	0	2
ERp18	O95881	172	EDEL	1	1
ERp44	Q9BS26	406	RDEL	0	1
ERp46	Q8NBS9	432	KDEL	0	3
TMX	Q9H3N1	280	Unknown	0	1
TMX2	Q9Y320	296	KKDK	2	1
TMX3	Q96JJ7	454	KKKD	2	1
TMX4	Q9H1E5	349	Unknown	1	1

Table 1: **Some sequence features of members of the human PDI family.** This table lists only members for which ER localization has been confirmed. In addition, Q96MT2 is probably a family member. The version of TMX2 included here represents the consensus version from human cDNA databases and is consistent with homologous proteins in a wide range of species. A longer version of TMX2 (Q8NBP9, 372 amino acids, of which the first 281 are identical to the sequence listed here) has been published (Meng, X., Zhang, C., Chen, J., Peng, S., Cao, Y., Ying, K., Xie, Y. and Mao, Y. (2003) Cloning and identification of a novel cDNA coding thioredoxin-related transmembrane protein 2. *Biochem. Genet.*, **41**, 99-106). Table adapted from Ellgaard, L. and Ruddock, L.W. (2005) The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep.*, **6**, 28-32.

No.	Buffer	PH	Salt	precipitation
1	Sodium acetate	pH 4.7	50 mM NaCl	++
2	Sodium acetate	pH 4.7	500 mM NaCl	+++
3	Sodium acetate	pH 5.3	50 mM NaCl	+++
4	Sodium acetate	pH 5.3	500 mM NaCl	++
5	Sodium citrate	pH 4.7	500 mM NaCl	+++
6	Sodium citrate	pH 5.5	500 mM NaCl	+++
7	Potassium phosphate	pH 5	50 mM KCl	+++
8	Potassium phosphate	pH 5	500 mM KCl	++
9	Potassium phosphate	pH 6	50 mM KCl	+++
10	Potassium phosphate	pH 6	500 mM KCl	+++
11	Potassium phosphate	pH 7	50 mM KCl	++
12	Potassium phosphate	pH 7	500 mM KCl	++
13	Sodium phosphate	pH 5.5	500 mM NaCl	+++
14	Sodium phosphate	pH 6.5	500 mM NaCl	+++
15	Sodium phosphate	pH 7.5	500 mM NaCl	+++
16	MOPS	pH 7	500 mM NaCl	+++
17	MOPS	pH 7.6	500 mM NaCl	++
18	HEPES	pH 7	50 mM NaCl	+++
19	HEPES	pH 7	500 mM NaCl	++
20	HEPES	pH 8	50 mM NaCl	+
21	HEPES	pH 8	500 mM NaCl	+
22	Tris-HCl	pH 8	500 mM NaCl	+
23	Tris-HCl	pH 9	500 mM NaCl	+
24	Ethanolamine	pH 9	500 mM NaCl	+++
25	Ethanolamine	pH 10	500 mM NaCl	+++
26	CAPS	pH 10	500 mM NaCl	++
27	CAPS	pH 11	500 mM NaCl	-
28	Sodium phosphate	pH 11.5	500 mM NaCl	+
29	Sodium phosphate	pH 12.5	500 mM NaCl	-

Table 2: **Precipitation of His-TMX4-Trx is least prominent in slightly basic Tris-HCl buffers.** Refolding screen performed on His-TMX4-Trx in different buffer conditions. Protein concentration was 2.33 mg/ml, the concentration of all buffers was 100 mM and all buffers contained 1 mM DTT. Very prominent and prominent precipitation (+++/++), precipitation less prominent but clearly detectable by light microscopy (+) and no precipitation (-).

No.	Buffer	PH	Salt, DTT	reagent	precipitation
1	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	5% glycerol	+++
2	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	400 mM sucrose	+++
3	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	1M L-arginine	-
4	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	2mM CHAPS	+++
5	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	2mM $MgCl_2$, 2mM $CaCl_2$	+++
6	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	-	+++
7	CAPS	pH 11	500 mM NaCl, 1mM DTT	-	+
8	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	2mM $CaCl_2$	+++
9	Tris-HCl	pH 8	500 mM NaCl, 1mM DTT	2mM $MgCl_2$	+++
10	Tris-HCl	pH 8	500 mM NaCl,	-	+++

Buffer 3 had to be pH-adjusted to 8, because I used Arginine-HCl

Table 3: **Precipitation of His-TMX4-Trx can only be prevented by 1 M L-arginine while the reagent is present in the solution.** Refolding screen performed on His-TMX4-Trx in 25 mM tris-HCl, pH 8, or 100 mM CAPS, pH 11, with addition of different reagents potentially promoting protein folding. Only the protein that was refolded in the presence of 1 M L-arginine was dialysed back to the original buffer free of L-arginine and was examined for precipitation. Very prominent and prominent precipitation (+++/++),precipitation less prominent but clearly detectable by light microscopy (+) and no precipitation (-).

Primer name	Primer sequence (5' → 3')
C5-11412fol-for*	AGG CAG CTG CGG CCG CAA CAT GGC GGG TGG GCG
C6-11412fol-rev	GCT CTA GAC TAC AGT CCC TTG TCA GC
C7-11412fol-HAfor*	GCT AGC TAC CCA TAC GAC GTC CCA GAC TAC GCT AGC GGC CCC GAG GAG GCC G
C8-11412fol-HArev*	GGG ACG TCG TAT GGG TAG CTA GCG GGG CCT GCC GTC GCC GC
3xmycfor	CGC TCG AGG AGC AAA AGC TCA TTT CTG AAG
3xmycrev	GCT CTA GAC TAA AGA TCC TCC TCG GAT ATT AAC
T4myc-Xhorev	-GCT GAC AAG GGA CTG CTC GAG GCT GCA TCT AGA GC
TMX4 C to S for	CAT GGT GTC CAT CCA GCC AGC AGA CTG ATT CAG
TMX4 C to S rev	TGA ATC AGT CTG CTG GCT GGA TGG ACA CCATG
TMX4 R to K for	GTGGTGGAAGACTCCTTGAAGCAGAAGAAAAGTCAGCATGCTGACAAG
TMX4 R to K rev	CTTGTCAGCATGCTGACTTTTCTTCTGCTTCAAGGAGTCTTCCACCAC
TMX4 lum+TM rev	GCT CTA GAC TAC TGC TCA GAA CGC TCA GAT AAA T
C4T41for	AGC TGC GGC CGC AAC ATG AAC CGG GGA GTC CC
C4T4F2rev	GCT CTA GAT CAA ATG GGG CTA CAT GTC TTC TG
C4T4F3rev	GAC ATA AGA CAT TGG CTG CAC CGG GG
C4T4F4for	CAG CCA ATG TCT TAT GTC TTT TTC GTC ATA G
C4T4F5rev	GAA ACA TTC TGA GAC ACA GAA GAA GAT GCC
C4T4F6for	CTT CTG TGT CTC AGA ATG TTT CTA TGT GCC
C4T4kozakfor	AGC TGC GGC CGC CAC AAT GAA CCG GGG AGT CCC
C4T4HAfor	5'GTG TAC CCA TAC GAC GTC CCA GAC TAC GCT AAG AAA GTG GTG CTG GG
C4T4HArev	GGG ACG TCG TAT GGG TAC ACC ACT TTC TTT CCC TGA GTG GC

Table 4: **Overview of sequences of oligonucleotide primers used in this study.**

A NON-CLASSICAL LOCALIZATION SIGNAL TARGETS THE NOVEL ER OXIDOREDUCTASE TMX4 TO THE ENDOPLASMIC RETICULUM

Doris Roth[1,2], Emily Lynes[3], Nils Althaus[2], Thomas Simmen[3] and Lars Ellgaard[1,*]

[1] Department of Molecular Biology, University of Copenhagen, August Krogh Building, 2100 Copenhagen, Denmark

[2] Institute of Biochemistry, ETH Zurich, 8093 Zurich, Switzerland

[3] Department of Cell Biology, University of Alberta, Canada

Running Title: A non-classical localization signal targets TMX4 to the ER

*Corresponding author. Tel. +45 35 32 17 25; Fax. +45 35 32 15 67; Email: lellgaard@aki.ku.dk

Key words: dpy-11, endoplasmic reticulum, protein disulfide isomerase, protein folding, localization motif, TMX4

ABSTRACT

To be added when everything else is complete

INTRODUCTION

The relatively oxidizing conditions inside the lumen of the endoplasmic reticulum (ER) provide an appropriate and supportive environment for oxidative protein folding (1). In the ER, the correct formation of disulfide bonds is often a prerequisite for the production of fully folded native proteins. During folding, pairs of cysteine thiols are oxidized to form covalent (disulfide) bonds. Since this process frequently leads to non-native combinations of cysteines, disulfide-bond reduction and isomerization (reshuffling) are critical reactions and often necessary steps en route to the correct structure (2).

Although disulfides form *in vitro* when oxygen is present, it is a very inefficient process in the absence of a catalyst. In the ER, oxidoreductases of the protein disulfide isomerase (PDI) family catalyze oxidation, isomerization and reduction by thiol-disulfide exchange. The PDI-family members contain at least one domain similar to thioredoxin, a cytosolic reductase of ~12 kDa. The three-dimensional structure of thioredoxin and similar proteins shows a typical α/β fold with a central β-sheet surrounded by α-helices. The residues of the CXXC active-site motif (where X denotes any amino acid residue) are located at the N-terminus of the second α-helix and in the preceeding loop.

In recent years, many novel genes that encode PDI-like proteins have been identified and a total of 17 such proteins are now known (3). The members of the human PDI family differ in size, the number and arrangement of thioredoxin-like domains, and in part in tissue distribution (3). Most of the redox-active PDIs share a glycine-histidine sequence for the two variable residues of the CXXC motif, but sequences using other residues exist. The nature of these two residues is known to influence the redox potential, i.e. the propensity of the active-site cysteines to be reduced or oxidized, and thereby the redox activity of a given enzyme (4,5).

For many human PDI-family members, the biological functions are unclear. Apart from possible differences in their redox activity, these proteins may well act on different substrates. For instance, ERp57 is likely to bind only certain cysteine-containing glycoproteins (6,7). The function of the PDI-family members is not restricted to oxidative folding. For example, ERp44 regulates Ca^{2+} release by the inositol 1,4,5-trisphosphate receptor 1, a Ca^{2+} channel of the ER membrane, through a

direct interaction that likely depends on the redox state of the receptor (8). It is also clear that some of the PDIs perform redox-unrelated functions since they lack both cysteines of the active site. An example is the murine ERp29 that is important for transport across the ER membrane to the cytosol of murine polyoma virus during infection (9). Here, in a step potentially required for membrane penetration, ERp29 promotes partial unfolding of the major structural viral protein VP1.

Recently, it became clear that while most of the PDI family members are soluble ER-luminal proteins, some span the ER membrane. This latter subset of the family is designated as the thioredoxin-like transmembrane – or "TMX" – proteins, and includes the TMX protein, TMX2, and TMX3. All contain a single thioredoxin-like domain but are otherwise not closely related. TMX has been shown to suppress apoptosis induced by brefeldin A when over-expressed in HEK293 cells (10). Whereas the knowledge about TMX2 is restricted to its cDNA sequence and tissue distribution (11), TMX3 is quite well studied *in vitro* in terms of structure-function relations (12). Nothing is known about substrate interactions of the TMX proteins but their membrane attachment could well influence substrate selectivity.

Here we characterize TMX4, a novel transmembrane member of the PDI family. We identify TMX4 as a paralog of TMX, and establish that it is an N-glycosylated type I transmembrane protein of the ER. TMX4 is expressed in a wide variety of human tissues. In living cells its unusual CPSC active-site sequence is kept in a predominantly oxidized state. Importantly, we show that in place of a classical ER-retention motif, an RQR sequence is required for efficient ER localization of the protein. The implications of the results are discussed in particular in relation to the closely related TMX protein.

EXPERIMENTAL PROCEDURES

Cell culture, Transfection and Antibodies – HeLa and Vero cells were maintained at 37 °C in α-Minimum Essential Medium (MEM) (Invitrogen) supplemented with 10% fetal calf serum (FCS) under 5 % CO_2. Vero cells and the HeLa cells used in Fig. 2 were transfected using an AMAXA electroporator following the manufacturer's recommendations for these cell lines. For all other experiments, HeLa cells were transfected with Lipofectamine 2000 (Invitrogen). The mouse monoclonal antibodies HA.11 (clone 16B12) against the HA epitope and clone 9E10 against the c-myc epitope were from Covance Research Products. The affinity purified rabbit antiserum specific for TMX3, α-TMX3-C, has previously been described (12). A. Helenius, ETH Zurich, generously provided the polyclonal rabbit antiserum against calnexin. Alexa Fluor 594 goat anti-mouse and Alexa Fluor 488 goat anti-rabbit were from Molecular Probes (Leiden, The Netherlands). Horseradish peroxidase-coupled goat anti-mouse and rabbit IgG antibodies were from Pierce.

Mammalian Expression Plasmids – The TMX4 cDNA clone 11412-f01 (IMAGE clone ID: 5165952) was obtained from MRC Geneservice (Cambridge, UK). It was amplified using the primers C5-11412fo1-for* and C8-11412fo1-HArev* and inserted into the pcDNA3.1 vector (Stratagene) using the restriction sites NotI and XbaI to yield the construct pcDNA3.1/TMX4. Table S1 in the Supplementary Material provides an overview of the sequences of all primers used in this study. All constructs were made by the cloning of PCR products into the pcDNA3.1 vector using the restriction sites NotI and XbaI, with the exception of pcDNA3.1/TMX4-myc where we used XhoI and XbaI. The correct sequences of all generated plasmids were verified by DNA sequencing.

The construct encoding N-terminally HA-tagged TMX4, pcDNA3.1/HA-TMX4, was generated by overlap extension PCR. Thus, the products from two initial separate PCR reactions (PCR(A) and PCR(B)) were used as templates in a third reaction (PCR(C)) together with the forward primer of reaction A and the reverse primer of reaction B. We used pcDNA3.1/TMX4 as a template for the two initial PCR reactions together with the primers C5-11412fo1-for* and C8-11412fo1-HArev* (PCR(A)) and C7-11412fo1-HAfor* and C6-11412fo1-rev (PCR(B)). The product of PCR(C) was inserted into pcDNA3.1 and the resulting construct, pcDNA3.1/HA-TMX4, was used as a template to generate pcDNA3.1/HA-TMX-KQK by means of the QuikChange Site-Directed Mutagenesis Kit (Stratagene) using the primers

"TMX4 R to K for" and "TMX4 R to K rev". As a result of the mutagenesis Arg338 and Arg340 were both replaced by lysines.

The pcDNA3.1/HA-TMX4-CPSS construct contains a CPSS active-site sequence instead of the endogenous CPSC. pcDNA3.1/HA-TMX4 was used as a template for PCR(A) and (B) with the primer pairs C5-11412fo1-for* and "TMX4 C to S rev" (PCR(A)) and C6-11412fo1-rev and "TMX4 C to S for" (PCR(B)). TMX4 devoid of the C-terminal tail, pcDNA3.1/Δtail, was generated using pcDNA3.1/HA-TMX4 as a template and the primers C5-11412fo1-for* and "TMX4 lum+TM rev".

For the construct encoding C-terminally myc-tagged TMX4, pcDNA3.1/TMX4-myc, we first equipped pcDNA3.1/TMX4 with a XhoI restriction sites before the stop codon in a PCR reaction with the primers C5-11412fo1-for* and T4myc-Xhorev. The resulting construct was named pcDNA3.1/TMX4xhoI. Then, a triple myc tag from the pcDNA3/Ero1β–3myc construct (kindly provided by R. Sitia, Milan) was amplified using the primers 3xmycfor and 3xmycrev. The PCR product was cloned into pcDNA3.1/TMX4xhoI using XhoI and XbaI restriction sites.

Human CD4 cDNA (13), kindly provided by H. P. Hauri, University of Basel (with the permission of K. Karjalainen), was subcloned into pcDNA3.1 after amplification using the primers C4T41for and C4T4F2rev to yield pcDNA3.1/CD4. Three chimeras between CD4 and TMX4 were then made by overlap extension PCR. PCR(A) was perfomed with pcDNA3.1/CD4 as template and C4T41for and C6-11412fo1-rev as primers. For PCR(B), pcDNA3.1/HA-TMX4 was used as template together with the primers C4T4F4for and C4T4F3rev. The product of PCR(C) was inserted into pcDNA3.1 to obtain pcDNA3.1/CD4-2. The construct pcDNA/CD4-3 was produced in a similar way, but using the primers C4T4F6for and C4T4F5rev in the PCR(B) reaction, and the primer pair C4T41for and C4T4F5rev for PCR(C). The plasmid pcDNA3.1/CD4-4 was generated as pcDNA/CD4-3, but with pcDNA3.1/HA-TMX4-KQK as a template in PCR(B).

The wild-type CD construct and the CD4-TMX4 chimeras were then HA tagged. PCR(A) was identical for all four, with pcDNA3.1/CD4 as template and the primers C4T4kozakfor and C4T4HArev. In PCR(B), C4T4HAfor was used as forward primer and combined with different templates and reverse primers: i) for pcDNA3.1/HA-CD4, the template was pcDNA3.1/CD4 and the primer C4T4F2rev, ii) for pcDNA3.1/HA-LTM$_{CD4}$C$_{TMX4}$, the template was pcDNA3.1/CD4-2 and the

primer C6-11412fo1-rev; iii) for pcDNA3.1/HA-$LTM_{CD4}C_{TMX4}$ the template was pcDNA3.1/CD4-3 and the primer C6-11412fo1-rev; iv) for pcDNA3.1/HA-$LTM_{CD4}C_{TMX4\text{-}KQK}$ the template was pcDNA3.1/CD4-4 and the primer C6-11412fo1-rev. For PCR(C), products A and B were combined with the forward primer C4T4kozakfor and the reverse primer corresponding to product B.

Immunofluorescence Microscopy – Transfected cells on microscope coverslips were fixed 24 h post-transfection using the method of Nakane (14). For studying intracellular proteins, cells were incubated for 4 h in fixing solution (10 mM sodium meta-periodate, 75 mM lysine, 375 mM sodium phosphate, pH 6.2, 2% paraformaldehyde), rinsed three times with PBS^{++} (phosphate-buffered saline; 154 mM NaCl, 1.9 mM KH_2PO_4, 8.1 mM Na_2HPO_4, pH 7.3, 1 mM $CaCl_2$, 0.5 mM $MgCl_2$) at room temperature, blocked with a solution of 3% bovine serum albumine (BSA) in PBS^{++} for 15 min and treated with permeabilization solution (0.05% saponin, 3% BSA in PBS^{++}) for 20 min. Next, cells were incubated with the monoclonal HA.11 antibody, diluted 1:1000 in permeabilization buffer, or the polyclonal α-CNX, diluted 1:500 in permeabilization buffer. After incubation for 1 h at room temperatur, the cells were washed three times in permeabilization buffer, followed by incubation for 45 min with secondary antibody, diluted 1:200 in permeabilization buffer. As secondary antibodies, we used Alexa 594 goat anti-mouse and Alexa 488 goat anti-rabbit. Coverslips were rinsed three times with permeabilization buffer, once with PBS^{++} and once briefly with water, and mounted on slides with Gel Mount™ Aqueous Mounting Medium (Sigma). After 1 h, the coverslip edges were resealed with Clarion™ Mounting Medium (Sigma). For immunoflourescence microscopy of cell surface exposed proteins on non-permeabilized cells, transfected cells were blocked and incubated with primary antibody on ice before fixation, followed by incubation with secondary antibody and mounting. Confocal images of fixed cells were acquired using either an inverted Zeiss LSM 510 metamicroscope (Carl Zeiss MicroImaging, Inc.) or a TCS NT/SP confocal laser scanning microscope (Leica Microsystems Heidelberg GmbH, Germany). Both microscopes were equipped with an apochrome 100X objective, NA 1.4.

Gel Electrophoresis and Western Blotting – Samples were separated on 10% 10x10.5 cm Hoefer minigels (GE Healthcare), with the exception of the CD4-TMX4 chimeras where a concentration of 7.5% polyacrylamide was used. For analysis of sucrose gradient fractions, SDS-PAGE was performed in a Hoefer standard vertical

unit (18x16 cm gels). Western blotting was performed as described in (12) using the ECL Advance Western blotting detection system (GE Healthcare). Antibodies were used in the following dilutions: HA.11 (anti-HA) 1:1000; α-TMX3-C (anti-TMX3) 1:500; horseradish peroxidase-coupled goat secondary antibodies (anti-mouse and rabbit) 1:50,000. Band intensities obtained by autoradiography or Western blotting were quantified using ImageJ (15).

Pulse-Chase Experiments and Immunoprecipitation – 12 h after transfection, cells were washed once with starvation medium (Dulbecco's modified Eagle's medium without methionine and cystine from Sigma), incubated in starvation medium for 15 min at 37 °C, and pulsed for 5 min or overnight at 37 °C with 125 μCi [^{35}S]methionine/cysteine (PerkinElmer) in 0.5 ml pulse medium per 35-mm dish. The cells were washed once with warm chase medium (α-MEM/10% FCS containing 5 mM methionine) and incubated in the same medium at 37 °C. At the end of the chase, dishes were transferred to ice where all subsequent steps were carried out. The cells were washed twice with PBS, scraped off with a rubber policeman in 1 ml lysis buffer (100 mM sodium phosphate, pH 8, 1% Triton X-100, and 1 mM phenylmethylsulfonyl fluoride (PMSF)), and passed 10 times through a 27 gauge needle. After incubation for 60 min, the lysate was centrifuged at 25,000 g for 30 min, and the supernatant was rotated for 2 h at 4 °C with 50 μl protein A-sepharose beads (GE Healthcare) preadsorbed with either 2 μl HA.11 or 2μl 9E10. After washing four times with lysis buffer and once with 100 mM sodium phosphate pH 8.0, proteins eluted from the beads by boiling in sample buffer .

Endoglycosidase H Digestion – Endoglycosidase H (Endo H) digestions were performed on HeLa cells lysed in loading buffer (180 mM Tris-HCl, pH 6.8, 0.5% SDS, and 1% 2-mercaptoethanol) and denatured at 96 °C for 15 min (Fig. 2), or on pulse-labeled immunoprecipitates after release from the beads by denaturation at 96 °C for 15 min in loading buffer (Fig. 7). Samples were then supplemented with one-tenth volume of 50 mM sodium citrate, pH 5.5, and treated with Endo H (New England Biolabs) for 1 h at 37 °C.

Membrane association and Proteinase K Protection Assay – To study the membrane association of TMX4, HeLa cells at 60% confluency were harvested, fractionated and extracted in sodium carbonate as previously described (12). To determine the topology of TMX4, HeLa cells transfected with constructs to express

HA-TMX4, TMX4-myc or ERp57-HA were labeled with [^{35}S]methionine/cysteine, washed twice with PBS, scraped off in homogenization buffer (20 mM Hepes-NaOH pH 7.5, 0.25 M sucrose, 1 mM dithiothreitol (DTT)) and passed 10 times through a 27 gauge needle. After removal of cell debris by centrifugation at 300 g for 3 min at 4 °C, the supernatant was centrifuged at 25,000 g for 1 h at 4 °C. The pellet was resuspended in assay buffer (25 mM Tris-HCl, pH 8, 500 mM NaCl, 1 mM DTT) and treated with 60 µg/ml proteinase K (Roche Molecular Biochemicals) on ice for 1 h in the presence or absence of 1% Triton X-100. After inhibition of proteinase K with 1 mM PMSF, detergent-free samples were solubilized with 1% Triton X-100 before the epitope-tagged proteins were immunoprecipitated, separated by SDS-PAGE and visualized by autoradiography.

Tissue Northern Blot – A human multiple tissue Northern blot (Ambion Inc) was hybridized with radioactive probes generated from a NotI-XhoI fragment obtained by restriction digestion of pcDNA3.1/TMX4-myc. The gel-purified fragment was labeled with [α-^{32}P]dCTP using the Strip-EZ DNA Kit (Ambion), followed by purification using a Micro Bio-Spin 30 column (Bio-Rad Laboratories) to remove unincorporated nucleotides. Hybridization was done at 65 °C in ULTRAhyb buffer (Ambion) following the manufacturer's protocol. The probed blot was visualized by autoradiography. The blot was then stripped according to the Strip-EZ DNA Kit protocol and hybridized with probes derived from the β-actin mouse DNA template delivered with the kit.

Determination of the in vivo redox state of TMX4 and TMX3 – Metabolically [^{35}S]methionine/cysteine-labeled HeLa cells transfected with pcDNA3.1/HA-TMX4 and pcDNA3.1/HA-TMX4-CPSS were modified in situ with N-ethylmaleimide (NEM), homogenized and lysed as described recently (16). Reduced and oxidized control samples generated by incubation of cells with 10 mM DTT and 5 mM diamide, respectively, were treated the same way. Immunoprecipitated HA-TMX4 and HA-TMX4-CPSS were used as the starting material for alkylation with 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS). Briefly, 27 µl of the sample were reduced with 1.5 µl 200 mM TCEP for 15 min and treated with 7.5 µl 75 mM AMS for 1h at room temperature. For the analysis of TMX3, cell lysates were used directly for alkylation by AMS (16). After separation by SDS-PAGE, TMX3

was visualized by Western blotting, and HA-TMX4 and HA-TMX4-CPSS by autoradiography.

Velocity Sedimentation on Sucrose Gradients – HeLa cells were lysed with 2% CHAPS and 1% Triton X-100 in HBS buffer (50 mM HEPES pH 7.5, 0.2 M NaCl). The lysates were then resolved on 5 ml 5-25% continuous sucrose gradiens in HBS supplemented with 0.1% Triton X-100, and centrifuged in a Sorvall AH-650 rotor at 42,000 rpm for 16 h at 4°C. The two standard proteins, BSA and β-amylase, were resolved in the same fashion. After the centrifugation, the gradients were fractionated manually from top to bottom. Fractions containing HA-TMX4 were analyzed by Western blotting, followed by quantification of band intensities as described above. The protein content in fractions containing standard proteins was determined using the Coomassie Plus Bradford Assay Kit (Pierce).

RESULTS

The TMX4 Sequence –We identified the human gene encoding TMX4 (EnsEMBL geneID: ENSG00000125827) in a GenBank BLAST search using a consensus sequence for a thioredoxin-like domain at the NCBI website (www.ncbi.nlm.nih.gov). The open reading frame encodes a protein of 349 residues (39 kDa, pI=4.3) including a signal sequence for entry into the secretory pathway with the signal peptidase cleavage site predicted between residues 23 and 24 (Fig. 1). The thioredoxin-like domain covers residues 39 to 136 and is followed by a stretch of 49 residues preceeding a predicted transmembrane helix spanning residues 188-210 (see also Fig. 3D). The C-terminal region (residues 211-349) is highly negatively charged with 54 acidic residues and contains no classical ER-localization motif of the KKXX type that targets many type I membrane proteins to the ER (17).

Phylogenetic analysis of the TMX4 and TMX sequences – Sequence alignments with other members of the human PDI family revealed that TMX4 is closely related to TMX. In particular, the thioredoxin-like domains show 53% sequence identity. In contrast, the C-terminal region beyond the predicted transmembrane helix is considerably less well conserved, with the most obvious difference being the length (78 residues in TMX versus 139 residues in TMX4). Transcripts for both proteins are present in a variety of mammals, amphibia, fish and birds but not in plants and yeasts, as shown by database analysis.

To learn more about the evolutionary relationship between the two proteins, we subjected a number of TMX and TMX4 orthologs to phylogenetic analysis using the maximum likelihood method. The evolutionary tree was rooted using ERp57, the closest homolog of PDI, and is shown in Fig. 1C. The proteins clustered in three main groups. The cluster shown at the bottom first diverged from a common ancestor and comprises organisms with only one TMX/TMX4 protein. One of these proteins is DPY-11 from *C. elegans* (18). Duplication then gave rise to the paralogs TMX and TMX4 that cluster in separate groups. Since the branch lengths represent the number of sequence changes that occurred prior to the next level of separation, the longer branches in the TMX4 cluster indicate that this protein diverged more from the common ancestor than TMX. Based on multiple sequence alignments (Fig. S1), we could pinpoint four residues in the thioredoxin-like domain that were defining for each protein. Using the NMR structure of the TMX thioredoxin-like domain and a structural model of TMX4 (Fig. S2) we analyzed potential functional consequences of

the substitutions at these four positions. This analysis is presented in the Supplementary Material.

TMX4 is an N-glycosylated transmembrane protein of the ER – To experimentally verify the predicted the ER localization of TMX4 we first generated a construct for expression of the N-terminally HA-tagged protein, HA-TMX4, by insertion of an HA epitope directly after the predicted signal peptidase cleavage site. Transfected Vero cells were fixed, permeabilized and incubated with primary antibodies against the ER protein calnexin (CNX) and the HA epitope. Cells were stained using fluorescently labeled secondary antibodies. By immunofluorescence microscopy, the exogenously expressed HA-TMX4 showed a typical reticular ER stain and was found to colocalize with CNX (Fig. 3A). [Similarly, endogenous TMX4 colocalized with CNX in A375P cells (Fig. 3X and see also Western blot analysis in Fig. 4B) ***IF data + Materials and Methods to be delivered by T. Simmen***].

The TMX4 sequence contains one consensus site for N-glycosylation on Asn46 (Fig. 1B). To investigate the glycosylation status of HA-TMX4 we expressed the protein in HeLa cells and treated the lysate with Endo H, a glycosidase that removes high-mannose sugars typically found on ER proteins. The subsequent Western blot analysis showed that HA-TMX4 indeed contained one glycan sensitive to Endo H treatment (Fig. 3B), as did endogenous TMX4 (Fig. 3C). [***Materials and Methods to be delivered by T. Simmen***].

The TMHMM algorithm (19) predicted TMX4 to be a type I transmembrane protein with residues 188-210 spanning the membrane (Fig. 3D). We next established that TMX4 is indeed a membrane protein (Fig. 3E). Isolated crude membranes from HeLa cells expressing HA-TMX4 were extracted with sodium carbonate at high pH, soluble and membrane-associated proteins were separated by high-speed centrifugation and probed by Western blotting. HA-TMX4 was detected exclusively in the membrane fraction, as was the case for the type I transmembrane protein CNX. The soluble marker protein ERp57 partly fractionated in the insoluble fraction as previously observed for another soluble ER protein, calreticulin, and possible due to the association with membrane proteins (12,20).

Membrane topology of TMX4 – In addition to the transmembrane domain, two other hydrophobic stretches were detected in TMX4 by the TMHMM algorithm (Fig. 3D). The first, at the very N-terminus, corresponded to the signal peptide and the second comprised residues 139-187, i.e. the region that links the thioredoxin-like

domain to the transmembrane domain. In principle, we could not rule out that this linker region was also buried in the membrane, thereby creating a topology with both termini of TMX4 in the ER and a short cytosolic loop. We therefore determined the membrane topology of TMX4 using a proteinase K protection assay (Fig. 3F).

Crude membrane extracts from [^{35}S]methionine/cysteine-labeled HeLa cells transfected with pcDNA3.1/HA-TMX4 were left untreated or treated with proteinase K, followed by immunoprecipitation of the expressed protein. The untreated sample showed a signal at the expected size for HA-TMX4 of ~45 kDa. Upon proteinase K treatment the band shifted to 25 kDa, a size that closely matched that of the 21 kDa expected for HA-TMX4 devoid of the C-terminal region. When the ER membrane was solubilized by addition of the detergent Triton X-100, proteinase K gained access to the lumen and the signal disappeared. So, the HA-tagged N-terminus was localized in the ER as expected based on the presence of the N-glycan on Asn46. To examine the orientation of the C-terminus, we performed the proteinase K assay on HeLa cells expressing C-terminally myc-tagged TMX4. Upon proteinase K treatment, the TMX4-myc signal disappeared almost completely, showing that the myc-tagged C-terminus was not protected inside the ER. The soluble lumenal ERp57 protein was used as a control for membrane integrity. Taken together, the results demonstrated that TMX4 is an N-glycosylated type I transmembrane protein of the ER.

Tissue distribution of TMX4 – To investigate the tissue distribution of TMX4, we first examined the transcript level in a variety of human tissues by Northern blotting (Fig. 4A). To this end, ^{32}P-labeled probes for TMX4 were used to hybridize a commercial tissue Northern blot. By autoradiography, we detected a signal at the predicted size of 6.0 kb for the TMX4 transcript (EnsEMBL transcript ID ENST00000246024) in all tissues tested, although with varying intensity. The highest transcript levels were found in brain, kidney, heart, skeletal muscle and lung.

[***Using our R504 antibody against TMX4 we also probed various human cell lines for the presence of the endogenous protein. We detected only low levels in common lines such as HeLa and HEK293, but a number of melanoma cells contained higher amounts of the protein (Fig. 4B). Awaiting final analysis by T. Simmen***]

Determination of the cellular redox state of TMX4 – The characterization of the *in vivo* redox state of oxidoreductases often gives first hints about their redox function. Here, we determined the oxidation state of ^{35}S-labeled HA-TMX4 expressed in HeLa cells. Our protocol involved a two-step alkylation process of thiols with the

alkylating agents NEM and AMS. First, we added the cell permeable NEM to living cells. After cell lysis, we treated lysates *in vitro* with the reducing agent TCEP and then AMS (see Experimental Procedures). Following these modifications, we immunoprecipitated HA-TMX4. By this method, NEM will block reduced cysteines in cellular proteins, whereas AMS will modify oxidized cysteines that have become available to alkylation as a results of TCEP treatment (16,21). Due to the larger size of AMS compared to NEM (536 Da versus 125 Da), oxidized proteins will show slower mobility by SDS-PAGE than reduced proteins. Cells treated with DTT and diamide were used as control samples for the reduced and oxidized state, respectively.

We detected two AMS-modified bands for HA-TMX4 (Fig. 5, lane 2). As expected, the lower band migrated like the reducing control (lane 3). Surprisingly though, the upper band did not run at the same height as the oxidizing control (lane 4), which showed an unexpectedly fast mobility. This was not caused specifically by diamide since another oxidizing reagent, dipyridyl disulfide, produced the same effect on TMX4 (data not shown). Moreover, the effect was specific for TMX4 as shown by using another PDI-family member, TMX3, for comparison (Fig. 5, lower panel). Like TMX4, this protein contains a single thioredoxin-like domain (12). In the experiment, we detected TMX3 by Western blotting in the exact same samples used in the upper panel for HA-TMX4. TMX3 showed the expected steady-state distribution of about 30% oxidized, 70% reduced (lane 6) (12,16). Importantly, the oxidized band ran at the height of the material in the diamide-treated controls (lanes 4 and 8) showing that diamide treatment was indeed effective.

TMX4 contains cysteines at positions 187, 213 and 326 in addition to the two found in the active site. We were able to ascertain that the upper band detected for HA-TMX4 at steady steate (lane 2, upper panel) was derived from the two cysteines in the active-site sequence (CPSC). Thus, a CPSS mutant that could not form the active-site disulfide bond showed only a single band at steady state (lane 6) migrating like the reduced control (lane 7). We concluded that a fraction of HA-TMX4 molecules at steady state have oxidized active-site cysteines. By densitometry we determined the oxidized fraction to be ~70%, with ~30% in the reduced state.

A non-classical motif is involved in ER targeting of TMX4 – Although TMX4 resides in the ER, it lacks a classical KKXX-type ER-localization signal. In multiple sequence alignments of TMX4 we noticed a LRQR motif conserved among species with only very few exceptions and located close to the C-terminus (Fig. 1B; Fig S1).

In a number of type II and multispanning membrane proteins a Φ/Ψ/R-R-X-R sequence, in which Φ/Ψ and X represent an aromatic/bulky hydrophobic and any residue, respectively, acts as an ER localization signal. So far, only one cellular type I membrane protein, the ER lectin called VIPL, has been shown to harbor an Arg-based ER-localization motif (22).

To test the possible involvement of the LRQR sequence in the ER localization of TMX4, we produced two HA-tagged mutants – a truncation variant lacking the C-terminal 121 residues (Δtail) and a mutant with the two arginines of the RQR sequence substituted by lysines (KQK) (Fig. 6A). The latter rather conservative mutation has been shown to efficiently abolish the function of the RXR sequence in Kir6.2, a subunit of K_{ATP} channels (23). After transfecting Vero cells, we analyzed the cellular localization of the wild-type protein and the two mutants by immunofluorescence microscopy (Fig. 6B). We found that, unlike the wild-type protein, HA-TMX4 lacking the cytosolic tail partially escaped the ER and reached the plasma membrane. The KQK mutant showed the same staining pattern as the Δtail mutant, being detected in the ER but also on the cell surface.

While some of the mutant TMX4 molecules trafficked to the cell surface, the majority of the three HA-TMX4 variants remained in the ER. To ascertain that this retention was not caused by aggregation as a result of over-expression, we used sucrose gradient velocity centrifugation. By this technique, aggregates sediment more slowly than monomeric or oligomeric proteins. HeLa cells over-expressing HA-TMX4 were solubilized in Triton X-100 and the lysates loaded onto a 5-25% continous sucrose gradient. After centrifugation, manually collected fractions were separated by SDS-PAGE and analyzed by Western blotting. For comparison, we ran in parallel samples of bovine serine albumin (BSA, a monomeric protein of 66 kDa that sediments at 4.3S) and β-amylase (a tetrameric complex of 200 kDa that sediments at 11.3S) (24) on the same type of gradients. The protein content of the fractions was determined by a Bradford assay. Quantification of the HA-TMX4 signals showed a discrete peak of the protein in fraction 3, ahead of BSA that peaked in fraction 4 (Fig. 6C). The distribution of HA-TMX4 was relatively broad so that the presence of a small amount of lower oligomeric complexes, for instance containing HA-TMX4 and a substrate, cannot be ruled out. Clearly though, no evidence was found for high-molecular weight complexes, showing that ER retention of HA-TMX4

was not an artifact of over-expression. We concluded that the RQR sequence in the C-terminal region contributes to the ER localization of HA-TMX4.

ER localization conferred by the TMX4 cytosolic tail depends on the RQR sequence - To further characterize the importance of the TMX4 LRQR sequence, we created three chimeric proteins (Fig. 7A). All were HA-tagged chimeras between the cytosolic tail of TMX4 and the luminal region of the plasma membrane protein CD4 (13). Two fusion proteins contained the wild-type sequence for the cytosolic tail of TMX4 but differed with respect to the origin of the transmembrane domain (TMX4 or CD4), whereas the last chimera contained the CD4 luminal and transmembrane domain fused to the KQK mutant of the TMX4 cytosolic tail.

The chimeras and wildtype HA-CD4 were transfected into HeLa cells that were then analyzed by confocal immunofluorescence microscopy (Fig. 7B). As expected, we found HA-CD4 strongly staining the plasma membrane. In contrast, the two CD4 chimeras fused to a wild-type TMX4 tail sequence did not reach the cell surface. The nature of the transmembrane domain – whether from CD4 or TMX4 – did not influence ER localization. Interestingly, the chimeric protein of HA-CD4 fused to the cytosolic tail of TMX4 containing the KQK mutation was clearly detected on the plasma membrane and showed a distribution similar to wildtype HA-CD4.

Having established that simply mutating the RQR sequence in the cytosolic tail of TMX4 to KQK was enough to redirect the chimeric protein to the plasma membrane, we investigated in more detail the trafficking of the various chimeras from the ER to the cell surface. This was done using a pulse-chase approach in combination with immunoprecipitation and Endo H digestion. The HA-tagged CD4 fusion proteins were expressed in HeLa cells and metabolically labeled with ^{35}S-methionine/cysteine. Following chase times of up to 12 h, the proteins were isolated by immunoprecipitation before they were either digested with Endo H or mock treated. Upon SDS-PAGE, the proteins were detected by autoradiography (Fig. 8A).

As depicted in Fig. 7A, the luminal region of CD4 contains two N-glycans. One of these becomes modified by complex glycosylation in the Golgi while the other remains in a high-mannose form (25). Therefore, cell surface CD4 is only partially Endo H resistant. This is most clearly seen in the upper panel of Fig. 8A, where – after chase times of 0.5, 5 and 12 h – two bands appeared upon Endo H digestion. The lower band migrated like the completely sensitive protein (compare to the 0 h chase

time + Endo H) whereas the band that appeared after 0.5 h chase time ran in between the glycosylated and the fully deglycosylated forms. This partially Endo H-resistant form (labeled "R") was used as a measure of CD4 on the plasma membrane. For the two CD4-TMX4 fusion proteins (middle and lower panel, Fig. 8A) the R band partially overlapped with a background band (labeled "*"). To correct for this overlap, we first quantified the two bands together and then subtracted the intensity of the background band obtained from the corresponding mock-treated lane. The results from four independent experiments were clear and consistent. In the representative experiment shown in Fig. 8 we found that Endo H resistance reached 65% for wildtype HA-CD4, whereas the construct with the wild-type TMX4 tail remained completely sensitive (Fig. 8B). Importantly, when the RQR sequence was mutated to KQK, the Endo H resistance of the fusion protein was restored to 40%. These results demonstrated that the cytosolic tail of TMX4 is sufficient for ER targeting and that the RQR sequence is required for this ability.

DISCUSSION

Here, we provide the first characterization of the novel transmembrane PDI-family member, TMX4. Using phylogenetic analysis we showed that TMX4 and TMX are paralogs that contain closely related N-terminal ER-luminal regions including a thioredoxin-like domain, but quite different C-terminal cytosolic tails. Our analysis revealed that an evolutionarily older version of the two proteins is the DPY-11 protein from *C. elegans*. In the worm, the protein is expressed exclusively in the hypodermis and is required for body and sensory organ morphogenesis (18). Curiously, DPY-11 does not harbor the RQR motif but a KKTK sequence at the C-terminus indicating a different ER localization mechanism than TMX4 and TMX. Also, human TMX could not rescue a *dpy-11* mutant in *C. elegans* (18).

Most of the TMX4 orthologs contain the active-site sequence CPSC, whereas all TMX proteins have a CPAC motif. These are unusual active-site sequences – for the PDIs, the CGHC sequence is the most common with 16 occurrences among 25 domains with a CXXC sequence, whereas only two other domains contain a CPXC sequence. Studies have shown that although the redox potential is responsive to the XX dipeptide sequence, it is not straightforward to predict changes in the redox potential and catalytic properties from mutations in the XX sequence (26,27). The enzymes with the common CGHC active-site motif generally seem to have moderate

redox potentials that are neither especially oxidizing nor reducing with values around –160 mV measured for TMX3, ERp57 and PDI (12,28-30). Here, it is noteworthy that we found the cellular redox state of TMX4 to be significantly more oxidized than that of the three CGHC-containing proteins mentioned above (12,16,21). This finding indicates that the redox activity of TMX4 – or the regulation of its in vivo redox state – differs from that of previously characterized family members.

The exact redox properties conferred to TMX4 and TMX by the CPXC motif are not known. For TMX, redox activity has been demonstrated *in vitro* (10,31), but it is unclear how the protein compares in such assays to well-characterized family members such as PDI and ERp57. Unfortunately, *in vitro* studies of basic enzymatic properties, redox potential measurement and structure determination of TMX4 were prevented by the low solubility of the protein. Moreover, while the recombinant thioredoxin-like domain was easily purified from *E. coli*, circular dichroism measurements showed that the protein had a low melting temperature of ~35°C with a fraction of the protein partially unfolded even at room temperature (data not shown). A construct containing the entire ER-luminal region including the hydrophobic linker region between the thioredoxin-like domain and the transmembrane domain – which in theory could have served a role in covering a potential hydrophobic patch on the thioredoxin-like domain – did not show improved solubility or stability properties. These findings might indicate that folding and/or solubility of the ER-luminal region of TMX4 depends on N-glycosylation or ER-specific chaperones.

Our analysis of the TMX4 and TMX sequences showed that they display only few distinguishing features in the luminal region (see Supplementary Material). While in particular the substitution of an arginine – conserved in many redox-active PDIs (32) – in TMX for a glycine in TMX4 might impair the redox activity of the latter, the large differences observed in the cytosolic tails are more likely to be determining with respect to cellular function.

Not surprisingly, the cytoplasmic tail of DPY-11 is required for the function of the protein (18). Here, we identified an Arg-based sequence motif, LRQR, conserved in the C-terminal cytosolic tail of TMX4 throughout species that functioned as an ER-localization signal. In HA-TMX4, we found that substitution of the two arginines by lysines resulted in surface expression of the protein (Fig. 6). That the LRQR sequence constitutes a bona fide ER-localization signal was further supported by the finding that the two arginines were required to confer ER localization of chimeras with the

reporter protein CD4 that normally localizes to the plasma membrane. The observation that a significant fraction of the HA-TMX4 RQR→KQK mutant remained in the ER can likely be explained by association of TMX4 with substrates or binding partners. Alternatively, forward transport might simply be inefficient due to the lack of an ER exit signal. We also observed that only 65% of wildtype CD4 molecules achieved Endo H resistance even after 12 hours chase (Fig. 8), as previously seen by others (33). Given that these experiments were performed in HeLa cells that do not normally express CD4, the relatively slow and inefficient ER export could well reflect the lack of an endogenous system to ensure effective delivery of the protein to the plasma membrane.

Like TMX4, TMX harbors a conserved Arg-based sequence motif (Fig. S1) that we speculate also serves as a localization signal. It is interesting to consider why TMX4 and TMX seem to use an Arg-based signal and not the classical C-terminal KKXX motif. It is clear that in many respects these two types of signals are different (34,35). Whereas the KKXX-like motifs have been established to function by COPI-mediated retrieval from the ERGIC and cis-Golgi to the ER (35,36), a potential role for COPI-coated vesicles in Arg-based ER localization is at present rather speculative and based on *in vitro* data only (34,35,37). Unlike the KKXX sequences, Arg-based signals seem to function via a regulated retention mechanism where masking of the signal, for instance by multimeric assembly or regulatory cytosolic interaction partners, allows ER exit (34,37). In the case of TMX4 (and TMX) masking of the Arg-based motif, potentially in response to specific cellular conditions, might therefore allow the relocalization of the protein.

In addition to the RQR motif, the tail of TMX4 (and to a lesser extent TMX) contains a long stretch of negatively charged residues comprising potential phosphorylation sites on serine residues (Fig. 1B). The combination of these two sequence features has been shown to constitute a potential binding site for PACS-2, a cytosolic sorting protein that uses COPI to target its substrate proteins to the ER dependent on their phosphorylation state (38,39). While speculative at present, it is conceivable that the cellular localization and thereby the function of TMX4 and TMX is modulated by binding of regulatory sorting factors that target different binding sites in the cytosolic tail. Concerning a potential functional connection, it is interesting to note that PACS-2 promotes staurosporin-induced apoptosis (39), while over-expression of TMX has been found to protect cells from brefeldin A-induced

apoptosis (10). We are currently pursuing the idea that the sorting motifs in TMX4 and TMX and a potential interaction with PACS-2 and other cytosolic factors connect the two proteins to a function in apoptosis.

REFERENCES

1. Ellgaard, L., and Helenius, A. (2003) *Nat. Rev. Mol. Cell Biol.* **4**(3), 181-191
2. Jansens, A., van Duijn, E., and Braakman, I. (2002) *Science* **298**(5602), 2401-2403
3. Ellgaard, L., and Ruddock, L. W. (2005) *EMBO Rep.* **6**(1), 28-32
4. Krause, G., Lundstrom, J., Barea, J. L., Pueyo de la Cuesta, C., and Holmgren, A. (1991) *J. Biol. Chem.* **266**(15), 9494-9500
5. Grauschopf, U., Winther, J. R., Korber, P., Zander, T., Dallinger, P., and Bardwell, J. C. (1995) *Cell* **83**(6), 947-955
6. Oliver, J. D., van der Wal, F. J., Bulleid, N. J., and High, S. (1997) *Science* **275**(5296), 86-88
7. Elliott, J. G., Oliver, J. D., and High, S. (1997) *J. Biol. Chem.* **272**(21), 13849-13855
8. Higo, T., Hattori, M., Nakamura, T., Natsume, T., Michikawa, T., and Mikoshiba, K. (2005) *Cell* **120**(1), 85-98
9. Magnuson, B., Rainey, E. K., Benjamin, T., Baryshev, M., Mkrtchian, S., and Tsai, B. (2005) *Mol. Cell* **20**(2), 289-300
10. Matsuo, Y., Akiyama, N., Nakamura, H., Yodoi, J., Noda, M., and Kizaka-Kondoh, S. (2001) *J. Biol. Chem.* **276**(13), 10032-10038
11. Meng, X., Zhang, C., Chen, J., Peng, S., Cao, Y., Ying, K., Xie, Y., and Mao, Y. (2003) *Biochem. Genet.* **41**(3-4), 99-106
12. Haugstetter, J., Blicher, T., and Ellgaard, L. (2005) *J. Biol. Chem.* **280**(9), 8371-8380
13. Maddon, P. J., Littman, D. R., Godfrey, M., Maddon, D. E., Chess, L., and Axel, R. (1985) *Cell* **42**(1), 93-104
14. McLean, I. W., and Nakane, P. K. (1974) *J. Histochem. Cytochem.* **22**(12), 1077-1083
15. Abramoff, M. D., Magelhaes, P. J., and Ram, S. J. (2004) *Biophotonics Int.* **11**(7), 36-42
16. Appenzeller-Herzog, C., and Ellgaard, L. (2007) *Antiox. Red. Signal.*, in press

17. Teasdale, R. D., and Jackson, M. R. (1996) *Annu. Rev. Cell Dev. Biol.* **12**, 27-54

18. Ko, F. C., and Chow, K. L. (2002) *Development* **129**(5), 1185-1194

19. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6**, 175-182

20. Pagani, M., Fabbri, M., Benedetti, C., Fassio, A., Pilati, S., Bulleid, N. J., Cabibbo, A., and Sitia, R. (2000) *J. Biol. Chem.* **275**(31), 23685-23692

21. Chakravarthi, S., and Bulleid, N. J. (2004) *J. Biol. Chem.* **279**(38), 39872-39879

22. Nufer, O., Mitrovic, S., and Hauri, H. P. (2003) *J. Biol. Chem.* **278**(18), 15886-15896

23. Zerangue, N., Schwappach, B., Jan, Y. N., and Jan, L. Y. (1999) *Neuron* **22**(3), 537-548

24. Johansson, E., Majka, J., and Burgers, P. M. (2001) *J. Biol. Chem.* **276**(47), 43824-43828

25. Harris, R. J., Chamow, S. M., Gregory, T. J., and Spellman, M. W. (1990) *Eur. J. Biochem.* **188**(2), 291-300

26. Chivers, P. T., Prehoda, K. E., and Raines, R. T. (1997) *Biochemistry* **36**(14), 4061-4066

27. Quan, S., Schneider, I., Pan, J., Hacht, A. V., and Bardwell, J. C. (2007) *J. Biol. Chem.*, In press

28. Frickel, E. M., Frei, P., Bouvier, M., Stafford, W. F., Helenius, A., Glockshuber, R., and Ellgaard, L. (2004) *J. Biol. Chem.* **279**(18), 18277-18287

29. Raturi, A., and Mutus, B. (2007) *Free Radic. Biol. Med.* **43**(1), 62-70

30. Darby, N. J., and Creighton, T. E. (1995) *Biochemistry* **34**(51), 16770-16780

31. Matsuo, Y., Nishinaka, Y., Suzuki, S., Kojima, M., Kizaka-Kondoh, S., Kondo, N., Son, A., Sakakura-Nishiyama, J., Yamaguchi, Y., Masutani, H., Ishii, Y., and Yodoi, J. (2004) *Arch. Biochem. Biophys.* **423**(1), 81-87

32. Lappi, A. K., Lensink, M. F., Alanen, H. I., Salo, K. E., Lobell, M., Juffer, A. H., and Ruddock, L. W. (2004) *J. Mol. Biol.* **335**(1), 283-295

33. Itin, C., Schindler, R., and Hauri, H. P. (1995) *J. Cell Biol.* **131**(1), 57-67

34. Michelsen, K., Yuan, H., and Schwappach, B. (2005) *EMBO Rep.* **6**(8), 717-722

35. Zerangue, N., Malan, M. J., Fried, S. R., Dazin, P. F., Jan, Y. N., Jan, L. Y., and Schwappach, B. (2001) *Proc. Natl. Acad. Sci. USA* **98**(5), 2431-2436

36. Eugster, A., Frigerio, G., Dale, M., and Duden, R. (2004) *Mol. Biol. Cell.* **15**(3), 1011-1023

37. Nufer, O., and Hauri, H. P. (2003) *Curr. Biol.* **13**(10), R391-393

38. Kottgen, M., Benzing, T., Simmen, T., Tauber, R., Buchholz, B., Feliciangeli, S., Huber, T. B., Schermer, B., Kramer-Zucker, A., Hopker, K., Simmen, K. C., Tschucke, C. C., Sandford, R., Kim, E., Thomas, G., and Walz, G. (2005) *EMBO J.* **24**(4), 705-716

39. Simmen, T., Aslan, J. E., Blagoveshchenskaya, A. D., Thomas, L., Wan, L., Xiang, Y., Feliciangeli, S. F., Hung, C. H., Crump, C. M., and Thomas, G. (2005) *EMBO J.* **24**(4), 717-729

40. Tochio, N., Koshiba, S., Inoue, M., Kigawa, T., and Yokoyama, S. (2005) PDB accession number: 1X5E, to be published

FOOTNOTES

We thank the Institute of Biochemistry, ETH Zurich, for the continued support and all members of the Ellgaard lab for critical reading of the manuscript. Constructs and antibodies were kindly supplied by Ari Helenius, Roberto Sitia, Hans Peter Hauri and Klaus Karjalainen. Funding obtained from the ETH Zurich, The Roche Research Foundation and Novo Nordisk Fonden is gratefully acknowledged.

[1] The abbreviations used are: AMS, 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid; BSA, bovine serine albumin; CNX, calnexin; DTT, dithio-1,4-threitol; endoglycosidase H, Endo H; endoplasmic reticulum, ER; Ero1, ER oxidase 1; FCS, fetal calf serum; GSH, reduced glutathione; GSSG, oxidized glutathione; MEM, minimum essential medium; NEM, *N*-ethylmaleimide; PBS, phosphate-buffered saline; PDI, protein disulfide isomerase; PMSF, phenylmethylsulfonyl fluoride; PVDF, polyvinylidene fluoride; redox, reduction-oxidation: TCEP, Tris(2-carboxyethyl)phosphine; TMX, thioredoxin-like transmembrane protein.

FIGURE LEGENDS

FIG. 1. **Sequence features and evolutionary conservation of TMX4.**

A, graphical overview of TMX4. The following sequence features are indicated: predicted signal peptide (SP), N-glycosylation consensus site (schematic drawing of a glycan), thioredoxin-like domain (Trx), active-site motif (CPSC) and predicted transmembrane domain (TM). B, multiple sequence alignment TMX4 from *Homo Sapiens* (Ensembl Gene ID ENSG00000125827), *Erinaceus europaeus* (Ensembl Gene ID ENSEEUG00000014678), *Canis familiaris* (Ensembl Gene ID ENSCAFG00000005883), *Ornithorhynchus anatinus* (Ensembl Gene ID ENSOANG00000008415) and human TMX (Ensembl Gene ID ENSG00000139921). Black boxes represent amino acid identities and gray boxes show amino acid similarities. The following features are indicated above the sequence: Signal sequence (striped rectangle), CPXC active-site sequence (double arrow), transmembrane region (checkered rectangle) and RQR motif (diamonds). Secondary structure elements of the TMX thioredoxin-like domain as determined by NMR spectroscopy are added below the TMX sequence, with cylinders and arrows representing α-helices and β-strands, respectively (see also Fig. S2) (40).

FIG. 2. **Phylogenetic tree for TMX and TMX4.**
Amino acid sequences of TMX and TMX4 from a variety of organisms were obtained from the Ensembl genome browser (release 45, June 2007). Full-length sequences were aligned on the ch.EMBNET.org server, using ClustalW XXL. The alignments were carried out using the blosum series matrix with an open gap penalty of 10 and an extended gap penalty of 0.05. Phylogenetic trees were constructed using PhyML and Phylip. Human ERp57 was selected as outgroup. The TMX4 cluster is highlighted in grey. Separate branches are labeled with the name of the individual organisms. Purple, TMX; black, TMX4; green, organisms with only one version of the protein.

FIG. 3. **TMX4 is an N-glycosylated type I transmembrane protein of the ER.**
A, HA-TMX4 colocalizes with the ER marker calnexin (CNX). Vero cells were transfected with pcDNA3.1/HA-TMX4, fixed and analyzed by immunofluorescence using antibodies against the HA epitope and CNX. B, HA-TMX4 is endoglycosidase H (Endo H) sensitive. Lysates from HeLa cells transiently expressing HA-TMX4 were mock treated or treated with Endo H and subjected to SDS-PAGE before Western blotting with an HA-specific antibody. C, Western blot analysis with the R504 antibody against endogenous TMX4 in A375P melanoma cells. Lysates were

treated with Endo H as indicated. The asterisk denotes a background band. D, Transmembrane (TM) topology prediction of the TMX4 sequence. The TMX4 amino acid sequence was analyzed by the transmembrane hidden Markov model (TMHMM) and the probability for the individual residues to be in a transmembrane helix were plotted versus the residue number. The stretch between residues 5 and 35 constitutes the predicted signal sequence. E, HA-TMX4 is an integral membrane protein. After subjecting crude membranes of HeLa cells expressing HA-TMX4 to alkali extraction, the soluble and insoluble fractions were separated by ultracentrifugation through a sucrose cushion. The distributions of HA-TMX4 and the control proteins CNX (a type I transmembrane protein) and ERp57 (a soluble ER protein) were visualized by Western blotting. F, TMX4 is a type I transmembrane protein. HeLa cells expressing HA-TMX4, TMX4-myc and ERp57-HA, respectively, were labeled with [^{35}S]methionine/cysteine. Crude membranes were isolated and mock treated or treated with Proteinase K in the presence or absence of 1% Triton X-100 (TX-100). Finally, samples were immunoprecipitated as indicated and the reactions were analyzed by autoradiography.

FIG. 4. **Tissue distribution of TMX4.**

A, The TMX4 mRNA is expressed in a wide array of tissues. A human tissue Northern blot was radioactively probed for the presence of TMX4 transcripts (predicted transcript length: 6.1 kb) and analyzed by autoradiography. Note a background band (*) in all tissues that migrates at the height of 28S ribosomal RNA (4.8 kb). As a control, the blot was stripped and labeled against β-actin (2.0 kb transcript length with an additional 1.8 kb isoform in muscle tissues). B, Western blot analysis of extracts of the human melanoma cell lines M2, A375M, A375P, LT5-1 and DX3 with the R504 antibody against endogenous TMX4. In the lower panel the actin signal is shown as a loading control.

FIG. 5. **HA-TMX4 is partially oxidized *in vivo*.**

HeLa cells transiently expressing wild-type HA-TMX4 (active-site sequence: CPSC; lanes 1-4) or the CPSS active-site mutant (lanes 5-8) were metabolically labeled, treated with 20 mM NEM, lysed and immunoprecipitated with an anti-HA antibody (autoradiogram shown in upper panel). Following reduction by TCEP, the lysates were mock treated or treated with AMS and then separated by SDS-PAGE. Cells

were pretreated with 10 mM DTT and 5 mM diamide to obtain controls for the fully reduced and oxidized forms, respectively. Lower panel: Western blot against TMX3 performed on the same lysate as the TMX4 immunoprecipitation. The positions of the oxidized and reduced forms of both proteins are indicated. The atypical mobility of TMX3 in lane 2 was caused by a gel artefact.

FIG. 6. **The cytoplasmic RQR motif contributes to ER localization of TMX4.** A, schematic representation of HA-TMX4 variants used for localization studies: full-length wild-type HA-TMX4 (WT), HA-TMX4 lacking the cytoplasmic region (Δtail), and HA-TMX4 with the two arginines of the RQR sequence substituted by lysines (KQK). Further abbreviations as in Fig. 1A. B, mutants lacking the RQR motif partially escape the ER and reach the plasma membrane. Vero cells were transfected with the three TMX4 constructs and stained with the anti-HA antibody after fixation and permeabilization (left). For detection of protein on the cell surface, cells were incubated with anti-HA antibody before fixation (right). IF: immunofluorescence. C, HA-TMX4 behaves as a monomer in velocity centrifugation. Postnuclear supernatants of HeLa cells expressing HA-TMX4 were supplemented with 0.1% Triton X-100, followed by velocity centrifugation on 5-25% continuous sucrose gradients at 42000 rpm for 16 h at 4°C. After harvesting the gradient from top to bottom, the fractions were analyzed by SDS-PAGE and Western blotting (upper panel). The HA-TMX4 signal was quantified by densitometry and plotted as the percentage of total (lower panel). As markers, bovine serum albumin (4.3S; 66 kDa) and 0-amylase (11.3S; 200 kDa) were analyzed on parallel gradients and the peak fraction was determined by a Bradford assay.

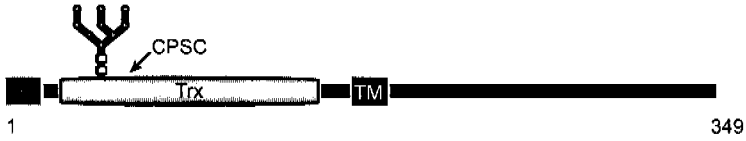
FIG. 7. **The TMX4 RQR sequence prevents cell surface expression of CD4.** A, schematic overview of HA-tagged CD4 and CD4/TMX4 chimeras. All possess the ER-luminal region of CD4 (L_{CD4}) but differ in the transmembrane domain (TM) and cytosolic tail (C), which are from CD4, TMX4 or TMX4-KQK as indicated. The N-glycosylation sites of CD4 are marked by asterisks (*). B, CD4 fusion proteins carrying the cytosolic region of WT-TMX4 do not reach the cell surface, whereas the cytosolic tail of TMX4-KQK is not sufficient to retain CD4 intracellularly. HeLa cells were transfected with the HA-CD4/TMX4 constructs, fixed and permeabilized or not

before staining with an anti-HA antibody to detect intracellular and cell surface expression, respectively, of the HA-CD4 chimeras.

FIG. 8. **The TMX4 RQR motif is required for ER localization of CD4.**
A, HeLa cells were transfected with HA-CD4 and chimeric HA-CD4/TMX4 constructs as indicated, metabolically labeled with [^{35}S]methionine/cysteine for 5 min and chased for the given periods of time. After immunoprecipitation, the HA-CD4/TMX4 variants were mock treated or treated with Endo H before the digest mixtures were separated by SDS-PAGE and visualized by autoradiography. Note that the band that represents the Endo H-resistant form (labeled 'R') migrated in between the undigested and completely sensitive forms. A background band that migrated slightly below the Endo H-resistant forms of HA-LTM$_{CD4}$-C$_{TMX4}$ and HA-LTM$_{CD4}$-C$_{TMX4\text{-}KQK}$ is labeled with an asterisk (*). B, Densitometric quantification of the autoradiograms in A. The percentage of Endo H-resistant protein was plotted against the chase time.

Figure 1

A


CPSC
Trx
TM
1
349

B

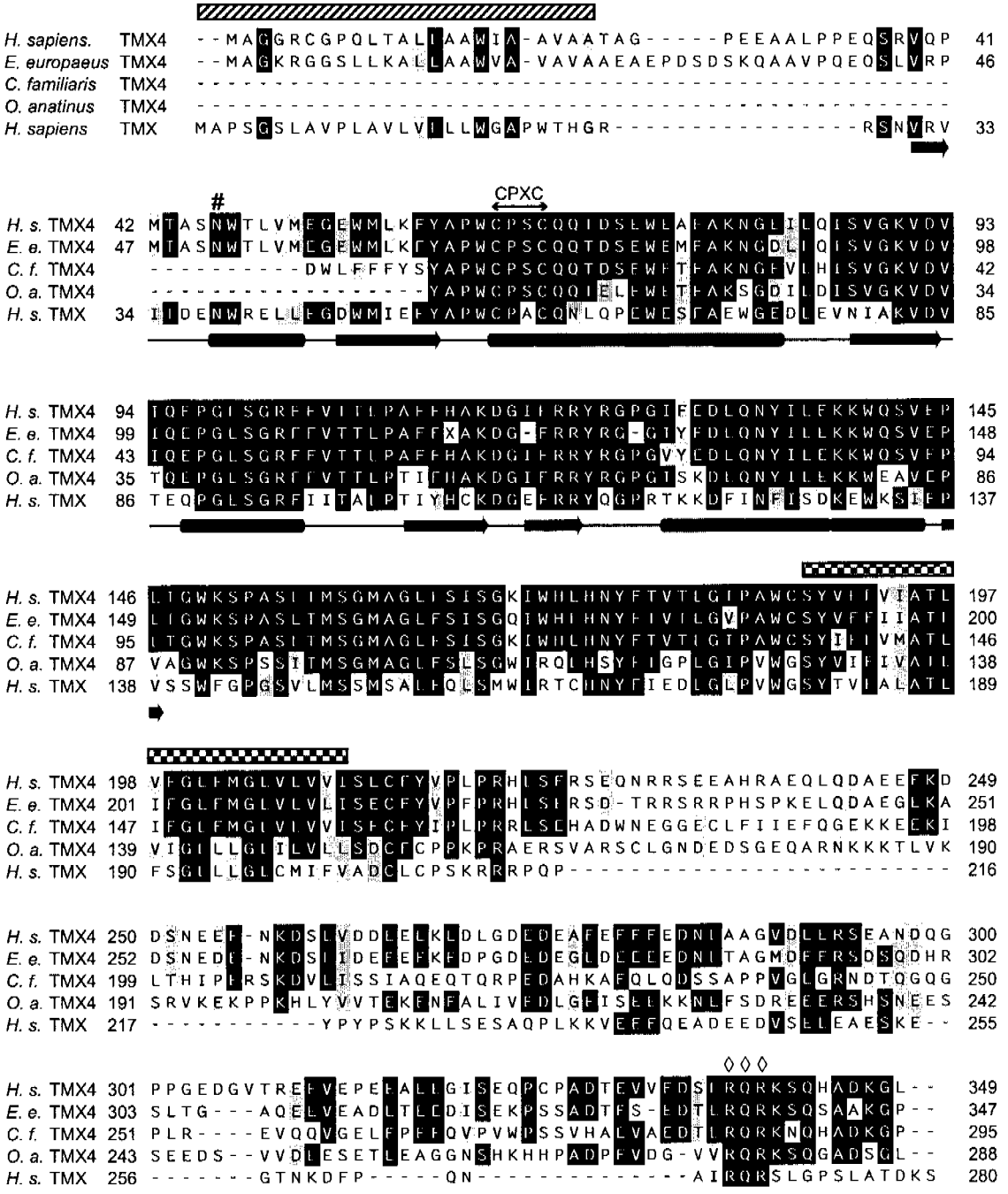
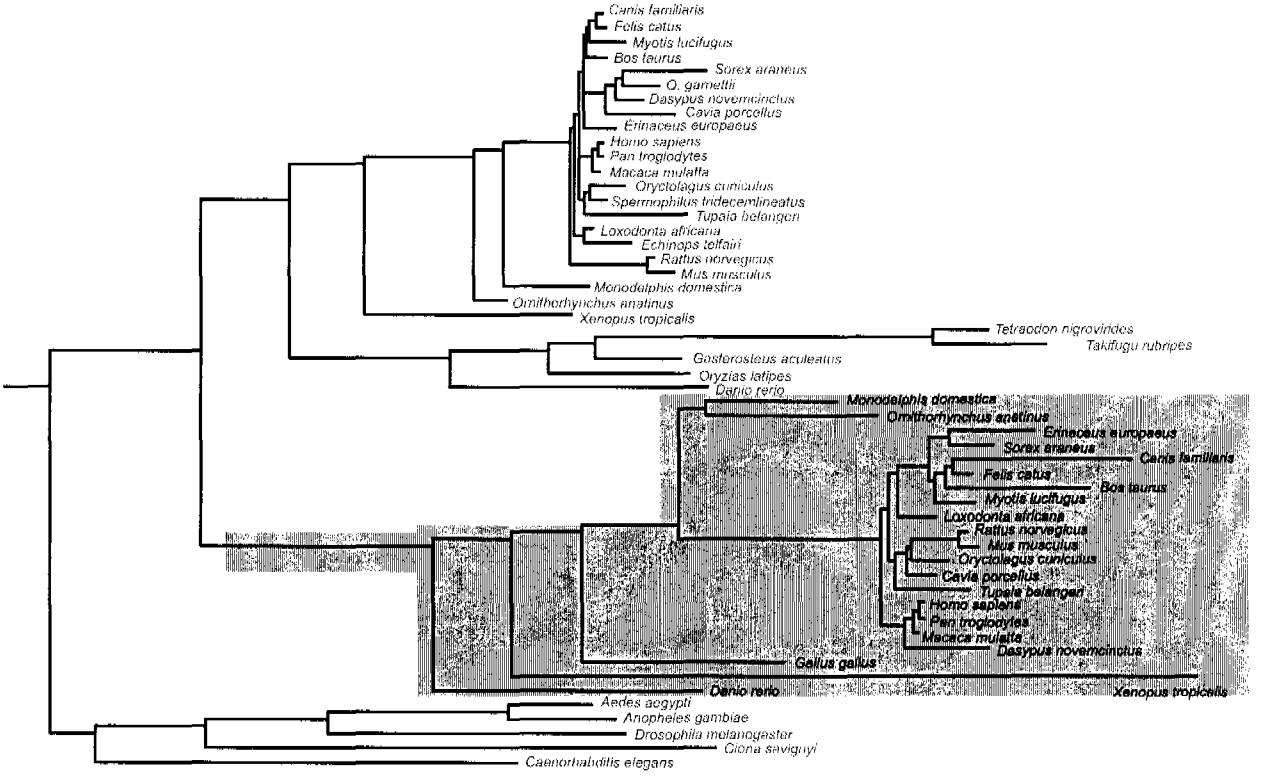
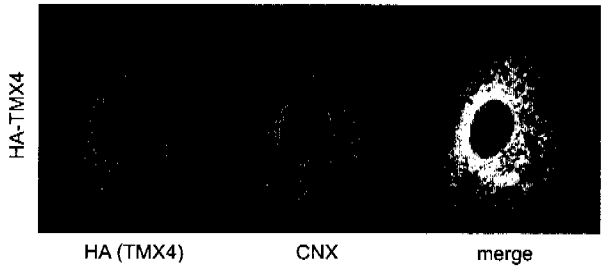

#
CPXC
H. sapiens. TMX4 41
E. europaeus TMX4 46
C. familiaris TMX4
O. anatinus TMX4
H. sapiens TMX 33
H. s. TMX4 42 93
E. e. TMX4 47 98
C. f. TMX4 42
O. a. TMX4 34
H. s. TMX 34 85
H. s. TMX4 94 145
E. e. TMX4 99 148
C. f. TMX4 43 94
O. a. TMX4 35 86
H. s. TMX 86 137
H. s. TMX4 146 197
E. e. TMX4 149 200
C. f. TMX4 95 146
O. a. TMX4 87 138
H. s. TMX 138 189
H. s. TMX4 198 249
E. e. TMX4 201 251
C. f. TMX4 147 198
O. a. TMX4 139 190
H. s. TMX 190 216
H. s. TMX4 250 300
E. e. TMX4 252 302
C. f. TMX4 199 250
O. a. TMX4 191 242
H. s. TMX 217 255
H. s. TMX4 301 349
E. e. TMX4 303 347
C. f. TMX4 251 295
O. a. TMX4 243 288
H. s. TMX 256 280

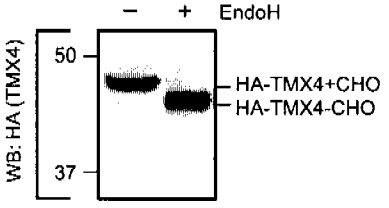
Figure 2


Canis familiaris
Felis catus
Myotis lucifugus
Bos taurus
Sorex araneus
O. garnettii
Dasypus novemcinctus
Cavia porcellus
Erinaceus europaeus
Homo sapiens
Pan troglodytes
Macaca mulatta
Oryctolagus cuniculus
Spermophilus tridecemlineatus
Tupaia belangeri
Loxodonta africana
Echinops telfairi
Rattus norvegicus
Mus musculus
Monodelphis domestica
Ornithorhynchus anatinus
Xenopus tropicalis
Tetraodon nigroviridis
Takifugu rubripes
Gasterosteus aculeatus
Oryzias latipes
Danio rerio
Monodelphis domestica
Ornithorhynchus anatinus
Erinaceus europaeus
Sorex araneus
Canis familiaris
Felis catus
Bos taurus
Myotis lucifugus
Loxodonta africana
Rattus norvegicus
Mus musculus
Oryctolagus cuniculus
Cavia porcellus
Tupaia belangeri
Homo sapiens
Pan troglodytes
Macaca mulatta
Dasypus novemcinctus
Gallus gallus
Danio rerio
Xenopus tropicalis
Aedes aegypti
Anopheles gambiae
Drosophila melanogaster
Ciona savignyi
Caenorhabditis elegans

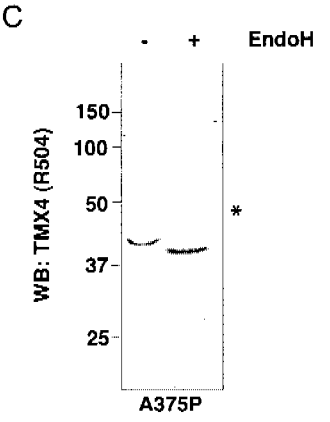
A


HA-TMX4
HA (TMX4)
CNX
merge

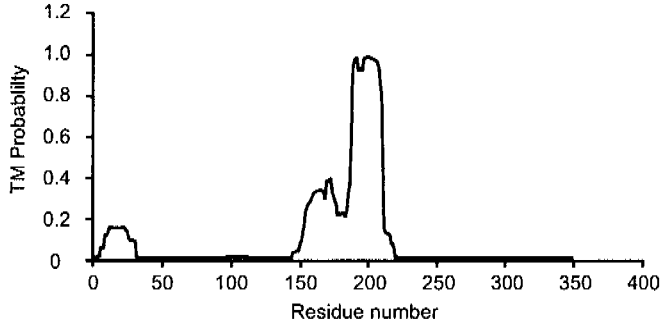
B


– + EndoH
WB: HA (TMX4)
50
37
HA-TMX4+CHO
HA-TMX4-CHO

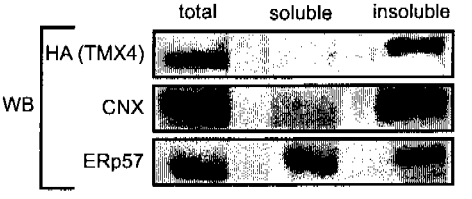
C


- + EndoH
WB: TMX4 (R504)
150
100
50
37
25
*
A375P

D


TM Probability
1.2
1.0
0.8
0.6
0.4
0.2
0
0 50 100 150 200 250 300 350 400
Residue number

E


total
soluble
insoluble
WB
HA (TMX4)
CNX
ERp57

F

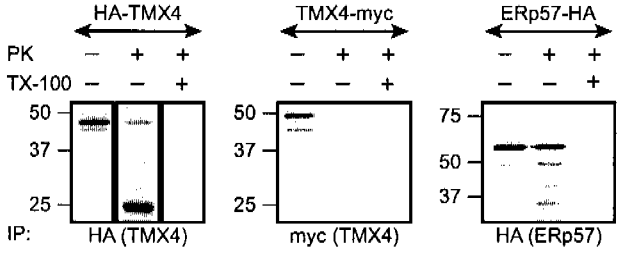
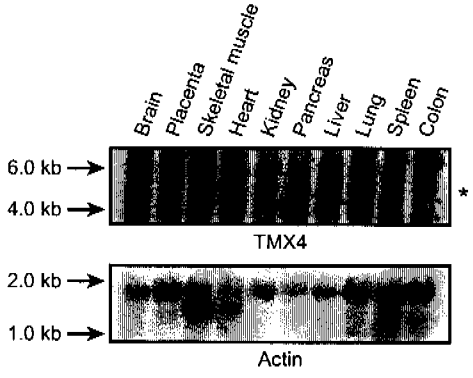
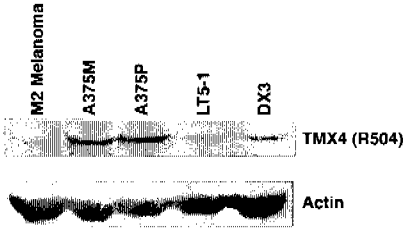

HA-TMX4
TMX4-myc
ERp57-HA
PK – + + – + + – + +
TX-100 – – + – – + – – +
50
37
25
75
50
37
IP:
HA (TMX4)
myc (TMX4)
HA (ERp57)

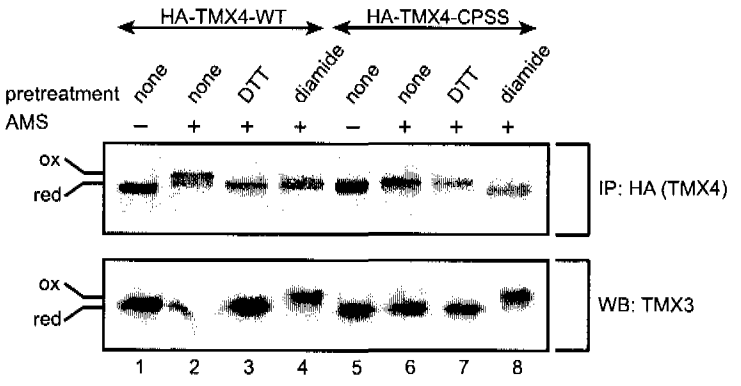
Figure 4

A

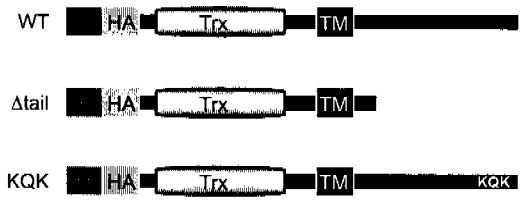

Brain
Placenta
Skeletal muscle
Heart
Kidney
Pancreas
Liver
Lung
Spleen
Colon
6.0 kb
4.0 kb
*
TMX4
2.0 kb
1.0 kb
Actin

B

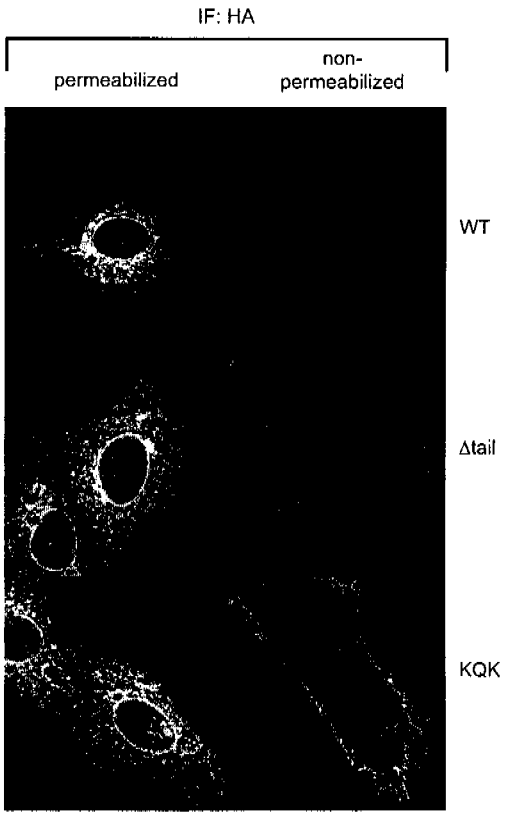

M2 Melanoma
A375M
A375P
LT5-1
DX3
TMX4 (R504)
Actin


HA-TMX4-WT
HA-TMX4-CPSS
pretreatment
AMS
none
none
DTT
diamide
none
none
DTT
diamide
– + + + – + + +
ox
red
IP: HA (TMX4)
ox
red
WB: TMX3
1 2 3 4 5 6 7 8

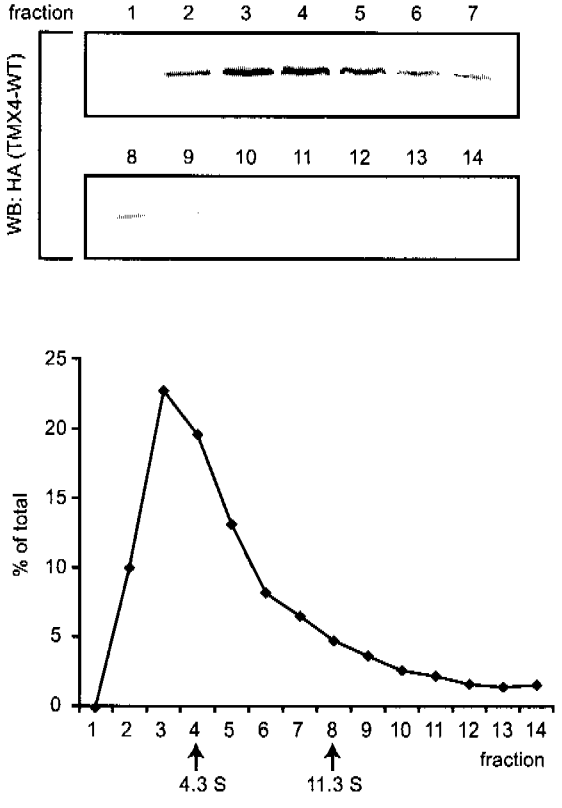
A


WT
HA
Trx
TM
Δtail
HA
Trx
TM
KQK
HA
Trx
TM
KQK

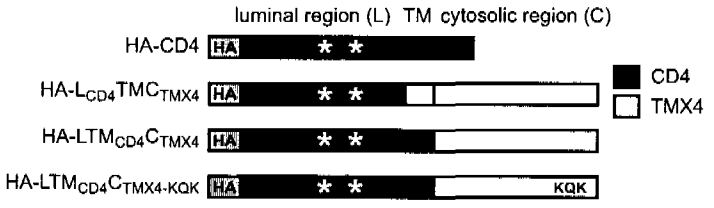
B


IF: HA
permeabilized
non-permeabilized
WT
Δtail
KQK

C


fraction 1 2 3 4 5 6 7
8 9 10 11 12 13 14
WB: HA (TMX4-WT)
25
20
15
10
5
0
% of total
1 2 3 4 5 6 7 8 9 10 11 12 13 14
fraction
4.3 S
11.3 S

A


luminal region (L) TM cytosolic region (C)
HA-CD4
HA-L$_{CD4}$TMC$_{TMX4}$
HA-LTM$_{CD4}$C$_{TMX4}$
HA-LTM$_{CD4}$C$_{TMX4-KQK}$
HA
KQK
CD4
TMX4

B

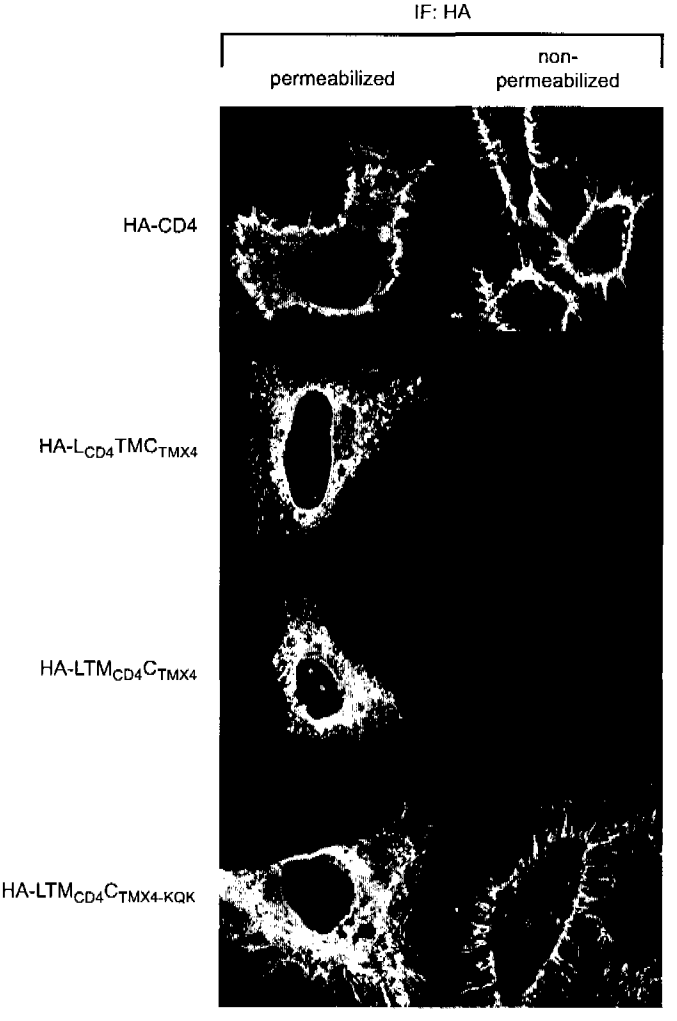
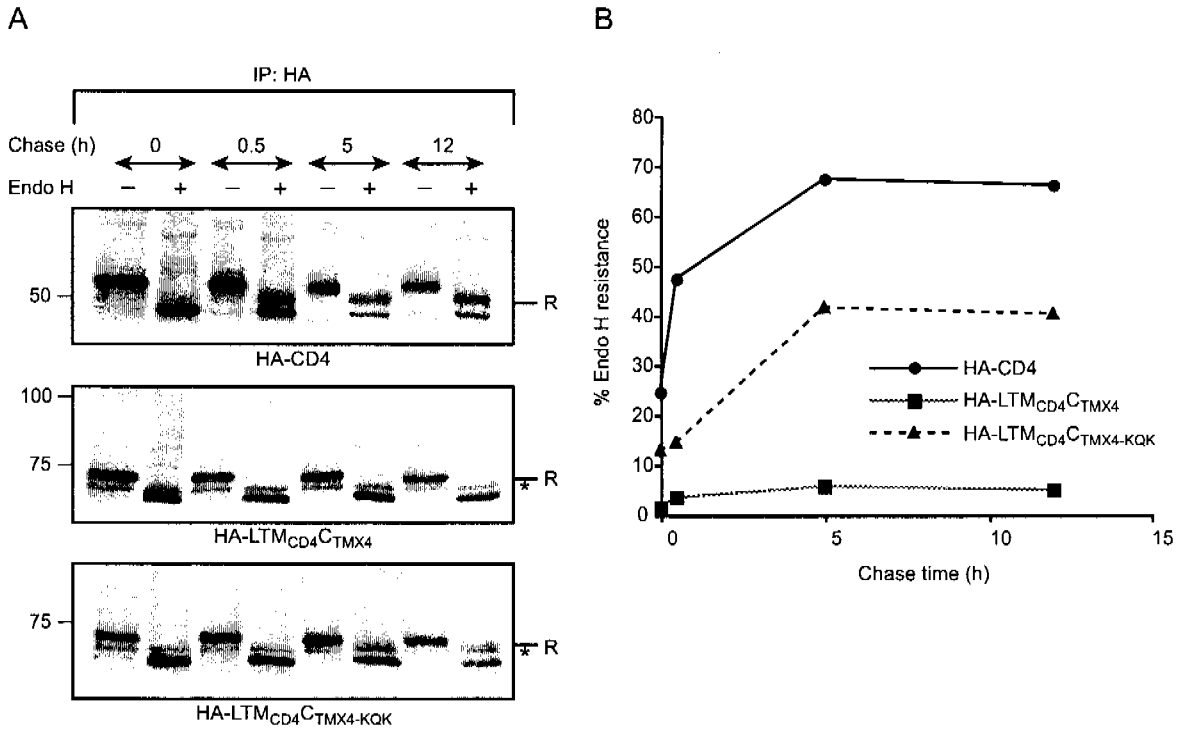

IF: HA
permeabilized
non-permeabilized
HA-CD4
HA-L$_{CD4}$TMC$_{TMX4}$
HA-LTM$_{CD4}$C$_{TMX4}$
HA-LTM$_{CD4}$C$_{TMX4-KQK}$

Figure 8


A
IP: HA
Chase (h)
0
0.5
5
12
Endo H
– + – + – + – +
50
R
HA-CD4
100
75
R
HA-LTM$_{CD4}$C$_{TMX4}$
75
R
HA-LTM$_{CD4}$C$_{TMX4-KQK}$
B
% Endo H resistance
80
70
60
50
40
30
20
10
0
0
5
10
15
Chase time (h)
HA-CD4
HA-LTM$_{CD4}$C$_{TMX4}$
HA-LTM$_{CD4}$C$_{TMX4-KQK}$