SELENOPROTEIN SEP15: FUNCTIONAL ANALYSIS AND ROLE IN QUALITY CONTROL IN THE ENDOPLASMIC RETICULUM

by

Vyacheslav Labunskyy

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Biochemistry

Under the Supervision of Professor Vadim N. Gladyshev

Lincoln, Nebraska

April, 2008

PREVIEW

UMI Number: 3297862

PREVIEW

UMI Microform 3297862

Copyright 2008 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

SELENOPROTEIN SEP15: FUNCTIONAL ANALYSIS AND ROLE IN QUALITY CONTROL IN THE ENDOPLASMIC RETICULUM

Vyacheslav Labunskyy, Ph. D.

University of Nebraska, 2008

Advisor: Vadim N. Gladyshev

Selenium is an essential trace element that is incorporated into proteins in the form of selenocysteine, the 21st natural amino acid which is encoded by the UGA codon. This rare amino acid is often present at the active centers of selenium-containing enzymes. One such protein, the 15-kDa selenoprotein (Sep15), has been recently reported to reside in the endoplasmic reticulum (ER) of eukaryotic cells and occur in a complex with UDP-glucose:glycoprotein glucosyltransferase (UGGT). UGGT is an ER chaperone and a component of the quality control machinery in the ER that assists in maturation of N-linked glycoproteins. Association of Sep15 with UGGT suggests that Sep15 may also be involved in assessing the structural fidelity of glycoproteins; however, the exact role of Sep15 in this process is poorly understood.

In this study, we characterized interactions between members of the Sep15 and UGGT protein families. We found that Sep15 and UGGT form a tight 1:1 complex, and their interaction is mediated by a cysteine-rich domain of Sep15. Mutation study of

PREVIEW

conserved cysteine residues showed that structural integrity of this domain is required for Sep15 to form a complex with UGGT.

Next, we employed nuclear magnetic resonance (NMR) to solve the structures of Sep15 and its distant homolog selenoprotein SelM. The solution NMR structures revealed that Sep15 and SelM have a thioredoxin-like fold. In mammals, Sep15 expression was regulated by dietary selenium, and both decreased and increased expression of this selenoprotein altered cellular redox homeostasis. The presence of thioredoxin-fold, together with the measured redox potential, suggested that Sep15 may catalyze reduction or isomerization of disulfide bonds in secreted proteins.

Finally, to address the role of Sep15 in protein folding, we analyzed whether expression of Sep15 is regulated by unfolded protein response in the ER. We found that Sep15 expression is affected by pharmacological ER stress, and its expression levels depend on the mechanism by which ER stress is induced. However, Sep15 deficiency did not itself result in accumulation of unfolded proteins suggesting that Sep15 may assist in maturation of a restricted group of N-linked glycoproteins and that other mechanisms may compensate for Sep15 function.

PREVIEW

This dissertation is dedicated to my son Daniel Marc and my wife Tanya

PREVIEW

ACKNOWLEDGEMENTS

I would like to express sincere gratitude to Dr. Vadim Gladyshev, my academic advisor, for giving me the opportunity to carry out my research project in his lab, as well as for his support and continuous guidance during my years as a graduate student. His enormous enthusiasm and optimism for science will be a continuous inspiration for me in the future career.

I would like to thank my supervisory committee members, Dr. Melanie Simpson, Dr. You Zhou, Dr. Marjorie Lou, and Dr. Donald Becker for their support, assistance and suggestions. I also wish to acknowledge my former committee member, Dr. Ruma Banerjee, for her advice and helpful discussions.

I also feel very fortunate to have friends and colleagues, whose support I have enjoyed throughout the years. Finally, I would like to thank my wife, Tanya, for her understanding and encouragement, my parents for their enormous support and my son Daniel.

PREVIEW

TABLE OF CONTENTS

PREVIEW

PREVIEW

ABBREVIATIONS

AMS, 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid

ATF6, activating transcription factor 6

BiP, binding protein

BrefA, brefeldin A

CNX, calnexin

CRT, calreticulin

Cumyl-OOH, cumyl hydroperoxide

DTT, dithiothreitol

eIF2α, eukaryotic initiation factor 2 alpha

ER, endoplasmic reticulum

ERp57, protein disulfide isomerase family A, member 3

Fep15, fish Sep15-like protein

GAPDH, glycelaldehyde-3-phosphate dehydrogenase

GSH, reduced glutathione

GSSG, oxidized glutathione

GPx1, glutathione peroxidase 1

GPx4, glutathione peroxidase 4

IRE1, inositol-requiring enzyme 1

MW, molecular weight

PREVIEW

NMR, nuclear magnetic resonance

PC, protein C

PDI, protein disulfide isomerase

PERK, PKR-like endoplasmic reticulum kinase

SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis

Se, selenium

Sec, selenocysteine

SECIS, selenocysteine insertion sequence

SelM, selenoprotein M

Sep15, 15-kDa selenoprotein

siRNA, small interfering RNA

SNP, single nucleotide polymorphism

t-BOOH, tert-butyl hydroperoxide

TR1, thioredoxin reductase 1

TR3, thioredoxin reductase 3

UGGT, UDP-glucose:glycoprotein glucosyltransferase

UPR, unfolded protein response

UPRE, unfolded protein response element

Xbp1, X-box binding protein 1

PREVIEW

CHAPTER 1

Introduction

PREVIEW

1.1. Selenoprotein Sep15

Selenium is an essential trace element that is present in 25 human proteins in the form of selenocysteine, the 21st natural amino acid which is encoded by the UGA codon (1, 2). Recent studies provide strong evidence that dietary selenium supplements reduce the incidence of cancer in animal models and human population (3-5). Low molecular weight selenium compounds and selenium-containing proteins have been implicated in this chemopreventive effect. However, the mechanism by which selenium suppresses tumor development remains to be established.

Several selenoproteins have been proposed as candidates that are responsible for the cancer preventive potential of selenium. One of these proteins is selenoprotein Sep15, which was identified several years ago in human T cells as a protein of unknown function (6). This selenoprotein was subsequently implicated in mediating the chemopreventive effect of selenium in certain types of cancers. Organs where selenium showed effect on cancer incidence include liver (7), prostate (8), breast (9), and lung (10).

PREVIEW

1.2. Potential Role of Sep15 in Cancer Prevention

The following lines of evidence make an argument for a possible role of Sep15 in cancer etiology. The human gene encoding Sep15 is located on chromosome 1 at position p31, a locus commonly deleted or mutated in human cancers and implicated in tumor suppression (11, 12). Sep15 expression was found to be significantly decreased in more than half of the tested cancer samples and tumor cell lines as compared to corresponding

controls (7, 10). Genetic analysis of the human Sep15 gene revealed the presence of two polymorphisms that show strong allelic association (7). These two single nucleotide polymorphisms (SNPs) are located at positions 811 (C/T) and 1125 (G/A) in the Sep15 mRNA (Fig. 1.1). The latter polymorphism is, in fact, part of the selenocysteine insertion sequence (SECIS) element, an RNA structure in the 3'-untranslated region required for recognition of in-frame UGA codons as selenocysteine. This SNP appears to influence expression of Sep15 in a selenium-dependent manner. The A1125 variant resulted in a higher expression of the selenoprotein, but it could not efficiently respond to the addition of selenium in the cell culture medium (7, 8, 13). Thus, this SNP directly influenced the expression levels of Sep15, and the outcome was dependent on the selenium status. These data suggested that individuals differing in these alleles may not only differ in Sep15 levels but also in the response to selenium dietary supplements. The polymorphisms have further relevance to cancer. The Al125 form is prevalent in African Americans who are known to have a higher incidence of prostate cancer (8). It is conceivable that African Americans require higher levels of selenium to achieve the protective levels of Sep15 expression. In addition, it was found that African Americans showed differences in allelic frequency in head, neck, and breast cancers and examples of the loss of heterozygosity at the Sep15 locus also were observed (8). Recent studies from other groups provide further evidence in support of these observations. For example, the A1125 form was found to be less responsive than the G1125 form to the selenium that showed growth inhibitory and apoptotic effects (10).

PREVIEW

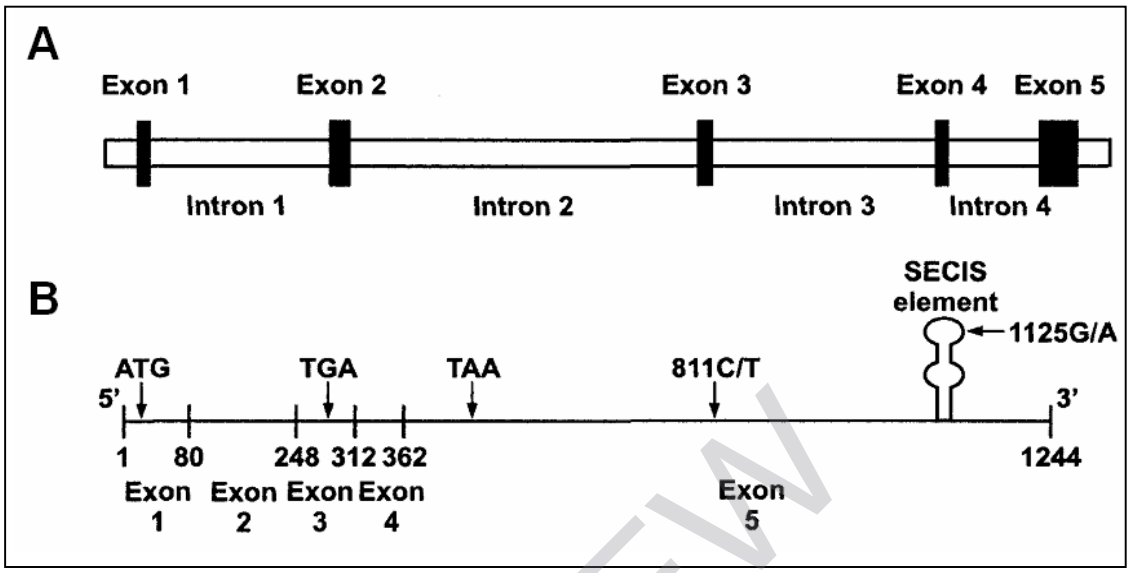


Figure 1.1. Structural organization of the human Sep15 gene. *A*, Exon-intron organization of the Sep15 gene. Closed squares correspond to exons, and horizontal lines correspond to introns and flanking regions. *B*, Organization of the human Sep15 cDNA sequence. The relative positions of the ATG initiation and the TGA Sec codons, the TAA termination signal and the single nucleotide polymorphisms (811C/T and 1125G/A) are indicated. The long horizontal line corresponds to the Sep15 cDNA, and short vertical lines correspond to exon-exon junctions. Numbers under junction sites correspond to last nucleotides in preceding exons.

PREVIEW

1.3. Expression Pattern and Regulation of Sep15 Expression by Dietary Selenium

Expression of the Sep15 gene was examined in several mouse and human tissues by northern blot and immunoblot analyses (7). The highest levels of gene expression were observed in prostate, liver, kidney, testes and brain, while lower levels were found in lung, spleen and skeletal muscle.

Dietary selenium has been shown to regulate the expression of selenoproteins and the abundance of corresponding mRNAs by acting at both transcriptional and translational levels. For example, selenium deficiency is known to result in up to a 99% decrease in the activity of glutathione peroxidase 1 (GPx1) and in up to a 90% decrease in the abundance of GPxl mRNA in livers of rats and mice. Similar to GPxl, Sep15 expression was also regulated by selenium availability in liver and kidney, although the changes in expression were less pronounced (14).

PREVIEW

1.4. The Sep15 Protein Family

Sep15 is highly conserved and is found in organisms from green algae to humans (6) suggesting an evolutionary conserved physiological function among Sep15 proteins. Recently, another ER resident eukaryotic selenoprotein, SelM, a distant homolog of Sep15, has been identified and characterized (15). Sep15 and SelM share 31% sequence identity in their proposed redox-active domains and form a distinct selenoprotein family. In mammals, SelM has a tissue expression pattern different from Sep15. Whereas the highest levels of Sep15 gene expression are observed in prostate, liver, kidney and testes, SelM is expressed predominantly in brain. Consistent with the ER localization, Sep15

and SelM encode N-terminal signal peptides, which are cleaved in the mature proteins (15, 16). Whereas SelM is likely retained in the ER by a C-terminal H/R/K-X-DL tetrapeptide, Sep15 lacks a typical ER retention signal suggesting that it is maintained in this cellular compartment by a different retention mechanism. Although the function of SelM has not been firmly established, recent studies provide evidence that decreased expression of this protein is associated with Alzheimer's disease and SelM may be involved in protection of neurons from oxidative damage (17).

More recently, a third member of the Sep15 protein family has been identified (18). This selenoprotein is found only in fish and was designated as Fep15 (for fish Sep15-like protein). Fep15 is also targeted to the ER by its N-terminal signal peptide and has a C-terminal ER retention signal (RDEL). Although Sep15, SelM and Fep15 share regions of significant sequence identity, they have several unique features that may specify differences in their substrate specificity and/or physiological functions. Multiple sequence alignment of the Sep15 family members revealed that SelM and Fep15 have an elongated C-terminus, whereas Sep15 possesses a distinct cysteine-rich N-terminal domain (Fig. 1.2). In contrast to SelM, in which the active-site cysteine and selenocysteine are organized into highly conserved C*xx*U motif (U is selenocysteine), Sep15 possesses an unusual motif, in which cysteine and selenocysteine are separated by only one amino acid (C*x*U). Moreover, Fep15 has only selenocysteine residue (U) and has valine in place of the conserved cysteine. Interestingly, Fep15 does not have any conserved cysteines, and in some of the Fep15 sequences cysteines are not present at all. Thus, this selenoprotein has a distinct catalytic mechanism that may involve formation of

PREVIEW

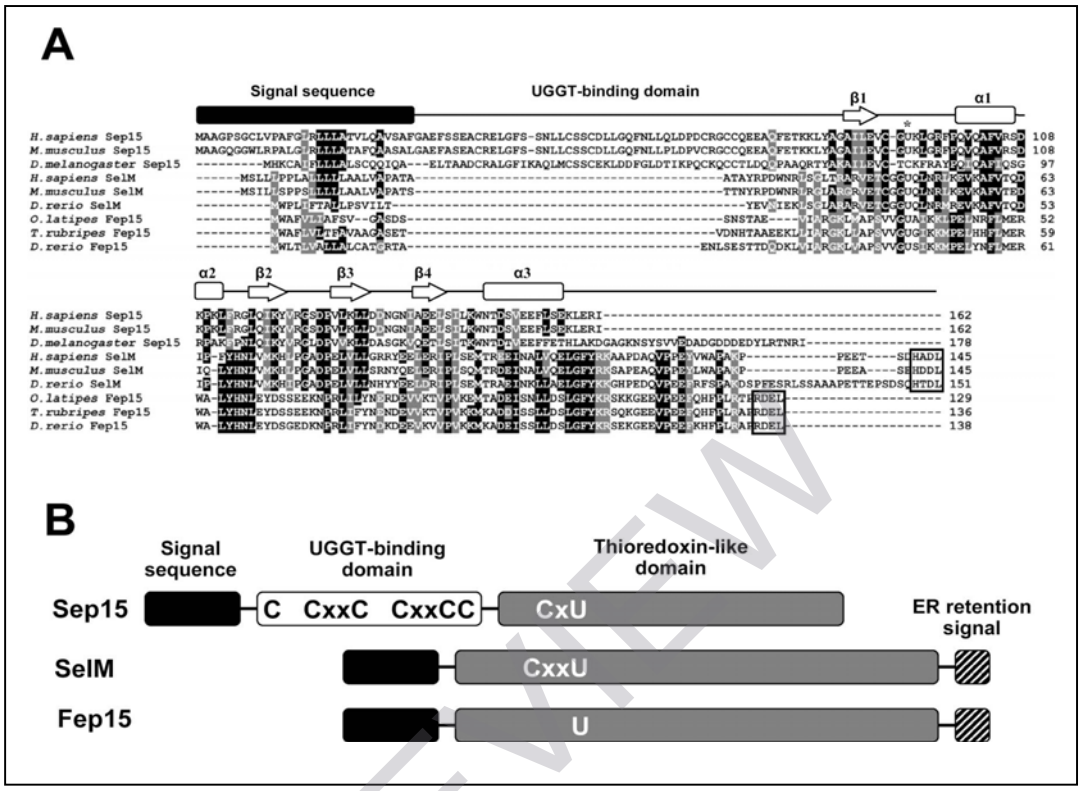


Figure 1.2. Alignment of the Sep15 family members. *A*, Multiple sequence alignment of Sep15, SelM and Fep15 proteins. Identical residues are shaded in black and similar residues in gray. The genetically encoded selenocysteine (*U*) residues and the corresponding cysteine residue found in the *D. melanogaster* Sep15 are marked with an asterisk. Predicted signal sequences and secondary structure elements are shown at the top. ER retention signals of SelM and Fep15 are highlighted with solid-line boxes. *B*, Schematic representation of domain arrangement of Sep15, SelM and Fep15. The Sep15 family proteins share an N-terminal signal sequence (colored *black*) and a common thioredoxin-like domain (colored *gray*). The cysteine-rich N-terminal extension of Sep15 has been labeled as the UGGT-binding domain.

PREVIEW

selenenic acid or selenenylsulfide bond with a cysteine residue in another protein or low molecular weight thiol, such as glutathione (19). Another unique feature of this novel selenoprotein is that it is present exclusively in selenocysteine-containing form, while other members of the Sep15 protein family exist in the form of both selenocysteine- and cysteine-containing proteins. Furthermore, the very narrow distribution of Fep15 among fishes suggests that this protein may have a highly specialized function.

1.5. The Calnexin Cycle

During initial purification from a human T-cell line, Sep15 was isolated in the denatured state (6); however, native Sep15 isolated from rat prostate and mouse liver co-purified with a protein of ~160-240 kDa. The binding partner of Sep15 was identified as UGGT, an ER chaperone and essential regulator of the calnexin (CNX) cycle (16). The CNX cycle is a quality control pathway localized to the ER that specifically assists in the folding of N-linked glycoproteins (20). In this quality control pathway, UGGT functions as the folding sensor that recognizes unfolded or improperly folded glycoproteins (Fig. 1.3). UGGT catalyzes the transfer of a glucose moiety from UDP-glucose to the glycan core (21) that creates a retention signal and initiates binding of membrane-bound CNX and its lumenal homolog calreticulin (CRT) to the glycan (22-25). This triggers the binding of ERp57 (a lumenal protein disulfide isomerase) to CNX and CRT, and accelerates folding by catalyzing disulfide bond exchange (26-30).

PREVIEW

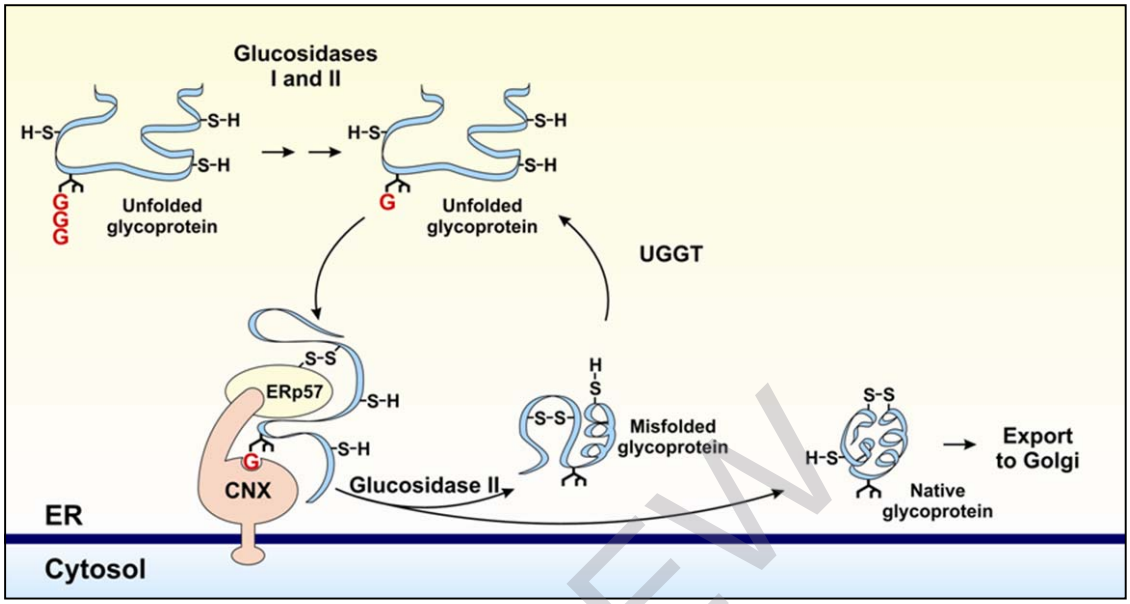


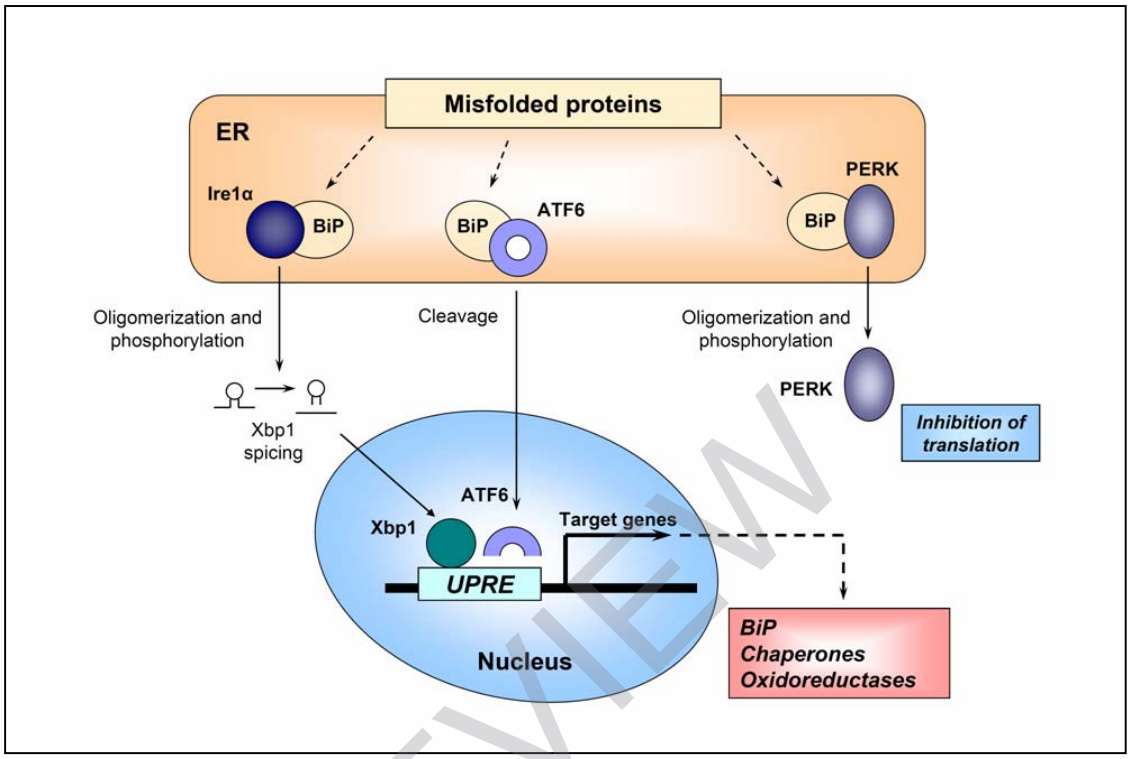
Figure 1.3. The CNX cycle. After removal of the two outermost glucose residues (G) from the N-linked oligosaccharide by glucosidases I and II, a newly synthesized glycoprotein binds to CNX or CRT (for simplicity only CNX is shown). ERp57, a lumenal protein disulfide isomerase associated with CNX (and CRT), catalyzes the formation of disulfide bonds in the unfolded glycoprotein substrate. When the remaining glucose residue is cleaved by glucosidase II, the glycoprotein dissociates. If the protein is properly folded, it is transported to Golgi. However, misfolded glycoprotein is reglucosylated by UGGT allowing it to reenter the CNX/CRT cycle.

PREVIEW

The loss of either CNX or CRT leads to perturbations in the folding of N-linked glycoproteins (31). While CRT deficiency has only mild consequences, the loss of CNX results in a dramatic impairment of folding, including decreased folding efficiency and retention of incompletely folded glycoproteins in the ER as well as to ER stress. In turn, accumulation of unfolded proteins in the ER leads to activation of unfolded protein response (UPR) (32). UPR is a signalling pathway that leads to inhibition of protein translation and enhanced expression of proteins that facilitate protein folding and help cells to remove incorrectly folded and/or excess proteins from the ER (Fig. 1.4). There are three proximal ER stress sensors: ATF6, IRE1 and PERK (33, 34) In unstressed cells, these proteins are associated with BiP protein, an ER-resident chaperone. But during ER stress, BiP binds misfolded proteins and releases these ER sensors. This leads to activation of ATF6 and Xbp1 transcription factors and increases expression of target genes: chaperones, oxidoreductases and BiP itself. In turn, released PERK phosphorylates initiation factor eIF2α and inhibits protein translation.

The role of Sep15 in protein folding and quality control processes is not known. However, the association of Sep15 and UGGT suggests that Sep15 may modulate the enzymatic activity of UGGT by functioning as a protein cofactor. Whether the loss of Sep15 also leads to accumulation of unfolded proteins in the ER, and whether Sep15 expression is up-regulated by UPR remains to be established.

PREVIEW



PREVIEW

Figure 1.4. The UPR signaling pathway. Three proximal ER stress sensors (IRE1, ATF6 and PERK) regulate the UPR. In unstressed cells, lumenal domains of these sensors are associated with BiP protein, an ER-resident chaperone. Upon accumulation of unfolded proteins in the ER, IRE1, ATF6 and PERK are released from BiP, which now binds unfolded proteins. These leads to activation of ATF6 and Xbp1 transcription factors and increases expression of target genes: chaperones, oxidoreductases as well as BiP. In turn, released PERK phosphorylates initiation factor eIF2α and inhibits protein translation. *UPRE*, unfolded protein response element.

CHAPTER 2

Interaction of Sep15 with UDP-glucose:Glycoprotein Glucosyltransferase

Note: The results described in this chapter have been published.

The authors and title of the paper are:

Labunskyy, V. M., Ferguson, A. D., Fomenko, D. E., Chelliah, Y., Hatfield, D. L., and Gladyshev, V. N. (2005) A Novel Cysteine-rich Domain of Sep15 Mediates the Interaction with UDP-glucose:Glycoprotein Glucosyltransferase. *J. Biol. Chem.* **280**, 37839-37845.

PREVIEW

2.1. Abstract

Although the precise function of Sep15 remains elusive, Sep15 copurifies with UDP-glucose:glycoprotein glucosyltransferase (UGGT), an essential regulator of quality control mechanisms within the ER. Two UGGT and two Sep15 homologs have been identified in mammals. In this study, we characterized interactions between these protein families. Sep15 and UGGT form a tight 1:1 complex, and these interactions are conserved between mammals and fruit flies. In mammalian cells, Sep15 co-immunoprecipitates with both UGGT isozymes. In contrast, a Sep15 homolog, selenoprotein SelM, does not form a complex with UGGT. Sequence analysis of members of the Sep15 family identified a novel N-terminal cysteine-rich domain in Sep15 that is absent in SelM. This domain contains six conserved cysteine residues that form two CxxC motifs that do not coordinate metal ions. If this domain is deleted or the cysteines are mutated, Sep15 no longer forms a complex with UGGT. Conversely, if the cysteine-rich domain of Sep15 is fused to the N-terminus of SelM, the resulting chimera is capable of binding UGGT. These data indicate that the cysteine-rich domain of Sep15 exclusively mediates protein-protein interactions with UGGT.

PREVIEW

2.2. Introduction

The precise biological function of Sep15 has not been firmly established; however, this selenoprotein co-purifies with the essential regulator of the CNX cycle designated UGGT (16). The CNX cycle is a quality control pathway localized to the ER that specifically assists in the folding of N-linked glycoproteins (20). Once a newly

synthesized protein enters the ER, a glycan core of 14 oligosaccharides ($Glc_3Man_9GlcNAc_2$) is covalently linked to appropriate asparagine residues. N-Linked glycoproteins are immediately recognized by glucosidase I and II, the first components of the CNX cycle, which sequentially trim the two outermost glucose moieties from the glycans core (35). Once processed, membrane-bound CNX and its lumenal homolog CRT bind to the glycan (22-25). This triggers the binding of ERp57 (a lumenal protein disulfide isomerase) to CNX and CRT and accelerates folding by catalyzing disulfide bond exchange (26-30). Regardless of the folded status of the bound glycoprotein, glucosidase II cleaves the remaining glucose residue and the complex disassociates, releasing the glycoprotein. If the glycoprotein is properly folded, it is rapidly exported from the ER by vesicular transport to the Golgi complex. Misfolded proteins are re-glucosylated by UGGT, a 175-kDa lumenal enzyme that catalyzes the transfer of a glucose moiety from UDP-glucose to the distal mannose of the glycan core (21). This event creates a retention signal that re-initiates the CNX cycle.

In this quality control pathway, UGGT functions as the folding sensor that assesses the folded status of glycoproteins. Although we do not understand how UGGT senses structural fidelity, several important clues have been revealed. (i) Sensing involves the innermost (and normally buried) N-GlcNAc glycosidic bond of the glycoprotein (36, 37). (ii) UGGT preferentially monoglucosylates partially folded late folding intermediates as compared with their fully denatured counterparts (38-40). (iii) UGGT is capable of distinguishing folded from misfolded domains within a single polypeptide (41, 42). (iv) Localized folding defects in otherwise correctly folded glycoproteins or

PREVIEW