

**Generation of
Endoplasmic Reticulum Protein 28 (ERp28)
Knock Out Mice, and
Structural and Functional Analysis of its
Drosophila Homologue, Wind**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Chaoshe Guo
aus Anyang, P.R. China

Göttingen 2003

D 7

Referent: Prof. Dr. H.D. Söling

Korreferent: Prof. Dr. H.J. Fritz

Tag der mündlichen Prüfung:

Acknowledgements

I thank Prof. Dr. Hans-Dieter Söling for his offer of this project, for his interest, support, invaluable criticisms, and critical reading of the manuscript. I am also grateful for his great help not only in science but also in other aspects of my stay in Göttingen.

To Prof. Dr. Hans-Joachim Fritz, of the Dept. of Molecular Genetics, University of Göttingen, my thanks for accepting the job as co-referee.

I am indebted to Prof. Dr. Reihard Jahn and colleagues, for providing the excellent working environment at the Department of Neurobiology, Max Planck Institute of Biophysical Chemistry.

Special thanks to Dr. David. M. Ferrari for his great help and guidance throughout the whole project including experimental planning, technique training, critical reading of the manuscript and most importantly, free discussions.

My warm thanks to Prof Dr. George. M. Sheldrick and Dr. Qingjun Ma at the Department of Structural Chemistry of the University of Göttingen, for their enthusiastic and successful collaborations in the crystallization and structural analysis of the protein, Wind. From them, I gained a lot of knowledge of X-ray crystallography.

Special thanks to Dr. Xunlei Zhou at the Department of Molecular Cell Biology, Max-Planck-Institute for Biophysical Chemistry, for taking the time to help me with the genomic DNA library screening, knock out vector construction and immuohistological staining.

I would like to thank Dr. Ahmed Mansouri and Mr. Sharif Mahsur for their providing ES cells, fibroblast feeder cells and for the training in ES cell culture, transfection and selection.

To Prof Dr. Nils Brose, at the Department of Neurobiology, Max-Planck Institute of Experimental Medicine, my thanks for his providing the mouse genomic DNA library.

I thank Mrs. Monika Schindler at the Max Planck Institute of Experimental Medicine for microinjection and thank the staff of the Animal Facility of the Max Planck Institute of Biophysical Chemistry for animal breeding.

Many warm thanks to my present and former colleagues, Mrs. Kathrin Barnewitz, Mrs. Sophie Verrier, Dr. Michael Wolde and others for their help in the last several years and for creating a very nice, friendly and productive working environment. Thanks also to Mrs. Kathrin Barnewitz for her reading of the manuscript.

Extra special thanks to my wife and my parents for their support and affection. And in addition, thanks to my friends Mrs. Ying Zhao, Dr. Xintao Zhang and Mr. Lingfen Luo and other friends for their help and support in the last few years.

Abbreviations

2-ME	β-mercaptoethanol
APS	ammonium persulphate
Bip	immunoglobulin heavy chain binding protein
BSA	bovine serum albumin
CaBP	calcium binding protein
CIP	calf intestinal alkaline phosphatase
CPY	carboxypeptidase Y
Cy3	indigocarboyanine
DAPI	4,6-diamidino-2-phenylindole
<i>D. discoideum</i>	<i>Dicyostelium discoideum</i>
<i>D. melanogaster</i>	<i>Drosophila melanogaster</i>
DEAE	diethylaminoethyl
DMEM	Dulbecco's minimal essential medium
DTT	dithiothreitol
D-V	dorsal-ventral
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum associated degradation
ERp	endoplasmic reticulum protein
ES cell	embryonic stem cell
EtOH	ethanol
FCS	foetal calf serum
Fig	Figure
GAPDH	glyceraldehydes phosphate dehydrogenase
GRp	glucose regulated protein
GSSH	glutathione (reduced)
GSSG	glutathione (oxidased)
GST	glutathione S-transferase
<i>H. sapiens</i>	<i>Homo sapiens</i>
HBS	Hepatitis B surface antigen
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
HSP	heat shock protein
Ig	Immunoglobulin
IPTG	isopropylthio-β-D-galactoside
Kb	Kilo base pairs
kDa	Kilo Dalton
KO	knock out
LDLR	low density lipoprotein-receptor
Lif	Leukemia inhibitor factor
<i>M. musculus</i>	<i>Mus musculus</i>
MHC	major histocompatibility complex
MOPS	2-[N-Morpholino]propanesulfonic acid
MTP	microsomal triglyceride transfer protein
MS	Mass spectrometry
NEM	N-ethyl maleimide
P4H	prolyl 4-hydroxylase
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline

Abbreviations

PCR	polymerase chain reaction
PDI	protein disulfide isomerase
PDI-D	PDI-related protein PDI-D
PDI-D α	PDI-related protein PDI-D α , redox active
PDI-D β	PDI-related protein PDI-D β , redox inactive
pfu	plaque forming unit
PPI	peptidyl prolyl cis-trans isomerase
PVDF	polyvinylidene difluoride
RAP	receptor-associated protein
<i>R. norvegicus</i>	<i>Rattus norvegicus</i>
RNase	ribonuclease
RT	room temperature
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	sodium dodecyl sulfate
TBE	Tri-borate-EDTA
TCA	trichloroacetic acid
TE	Tris-EDTA
TEMED	1,2-Bis-(dimethylamino)-ethane
Tris	tris[hydroxymethyl]aminomethane
TRX	thioredoxin
UPR	unfolded protein response
UPRE	unfolded protein response element
UV	Ultraviolet
v/v	volume per volume
w/v	weight per volume
wt	wild type
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactoside

Amino acids and their symbols

Three-letter symbols	One-letter symbols	Amino acids
Ala	A	Alanine
Cys	C	Cysteine
Asp	D	Aspartic acid
Glu	E	Glutamic acid
Phe	F	Phenylalanine
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Lys	K	Lysine
Leu	L	Leucine
Met	M	Methionine
Asn	N	Asparagine
Pro	P	Proline
Gln	Q	Glutamine
Arg	R	Arginine
Ser	S	Serine
Thr	T	Threonine
Val	V	Valine
Trp	W	Tryptophan
Tyr	Y	Tyrosine

Contents

1. Summary	1
2. Introduction	3
2.1 Endoplasmic reticulum (ER) and protein folding	3
2.2 Protein disulfide Isomerase	4
2.2.1 Redox / Isomerase function of PDI and the pathway of formation of disulfide bonds	5
2.2.2 Chaperone function	8
2.2.3 Other functions	9
2.2.3.1 Subunit association	9
2.2.3.2 Different Localization of PDI	9
2.2.4 The structure of PDI and its functions	10
2.2.4.1 Structural comparison between thioredoxin and a and b domains of PDI	11
2.2.4.2 The importance of the b' domain of PDI	12
2.3 PDI family or PDI-L (like) family	13
2.3.1 ERp57	15
2.3.2 PDIp	15
2.3.3 ERp72 (CaBP2)	16
2.3.4 PDIr	16
2.3.5 ERdj5 (JPDI)	16
2.3.6 ERp44	17
2.3.7 ERp18	17
2.3.8 Calsequestrin (CSQ)	18
2.3.9 P5 (CaBP1)	19
2.4 PDI-D subfamily	19
2.4.1 ERp28	21
2.4.2 Wind and the Dorsal-Ventral patterning of the <i>Drosophila</i> embryo	21
3. Aims of the work	25
4. Materials	26

5. Methods	28
5.1 General methods	28
5.1.1 PCR (polymerase chain reaction)	28
5.1.2 Purification of PCR products	29
5.1.3 DNA extraction from agarose gel	29
5.1.4 Estimation of DNA purity and quantitation	29
5.1.5 Plasmid extraction	29
5.1.6 Restriction enzyme digestion, dephosphorylation, preparation of blunt-ended DNA and ligation	29
5.1.7. Transformation of <i>E. coli</i>	30
5.1.7.1 Preparation of competent <i>E. coli</i> for heat shock transformation	30
5.1.7.2 Preparation of electro-competent <i>E. coli</i>	31
5.1.7.3 Transformation by heat shock	31
5.1.7.4 Transformation by electroporation	31
5.1.8 DNA gel electrophoresis	32
5.1.9 Protein gel electrophoresis	32
5.1.9.1 Polyacrylamide gel electrophoresis (PAGE)	32
5.1.9.2 SDS-PAGE gel	32
5.1.9.3 Native gel (non-denaturing gel)	34
5.1.10 Coomassie brilliant blue staining	34
5.1.11. Western blotting	35
5.1.11.1 Transferring of proteins onto nitrocellulose membranes	35
5.1.11.2 Antibody staining of Western blots	35
5.1.11.3 Detection with peroxidase/diamino benzidine	36
5.1.11.4 Chemiluminescence	36
5.1.12 Protein quantitation	36
5.2 Generation of knock out mice	37
5.2.1 A brief introduction of the knock out strategy and the knock out vector	37
5.2.2 Searching and alignment of the mouse ERp28 ESTs from the database	38
5.2.3 The exon distribution of mouse ERp28	38
5.2.4 RNA extraction	39
5.2.5 RT-PCR to obtain a cDNA probe for mouse Genomic DNA library screening	

5.2.6 TA cloning	40
5.2.7 Genomic DNA library screening	40
5.2.7.1 Library titre determination	40
5.2.7.2 The first round of screening	40
5.2.7.3 The second round of screening	43
5.2.7.4 The third round of screening by PCR	43
5.2.8 Isolation of DNA from the positive plaques	44
5.2.9 Restriction mapping	44
5.2.10 Steps for building the knock-out vector	44
5.2.11 Linearisation of the knock out vector	44
5.2.12 ES cell transfection and selection	44
5.2.13 Extraction of genomic DNA from the surviving ES cells	50
5.2.14 Screening of the recombinant ES cells by PCR	50
5.2.15 Preparation of positive ES cells for blastocyst injection	51
5.2.16 Generation of chimeric mice by microinjection	51
5.2.17 Preparation of genomic DNA from the tails of chimeric mice and genotyping by PCR	52
5.2.18 Crossing of the chimeric mice with wild type mice to generate heterozygotes	52
5.2.19 Preparation of genomic DNA from the tails of possible heterozygotic mice and verification by PCR	52
5.2.20 Recrossing of male and female heterozygotes to produce homozygotes	52
5.2.21 Genotyping of chimera, heterozygotes and homozygotes by Southern blotting	52
5.2.22 Immunohistochemistry on paraffin sections	54
5.3 Methods for structural and functional analysis of Wind	56
5.3.1 Construction of His-Wind and Wind-His expression vectors	56
5.3.2 Expression of His-Wind and Wind-His in XL1-Blue	56
5.3.3 Purification of His-Wind and Wind-His	56
5.3.4 Dialysis and concentration	57
5.3.5 Immunisation of rabbits to generate antibodies against His-Wind	58
5.3.6 IgG purification using Protein A Sepharose	58
5.3.7 Removal from anti-Wind antiserum of antibodies against <i>E.coli</i> protein using CNBr activated Sepharose 4B	58

5.3.8 Insulin reduction assay	59
5.3.9 Cross-linking of His-Wind/Wind-His	60
5.3.10 Structure determination of Wind	60
5.3.10.1 Crystallization	60
5.3.10.2 Heavy atom derivative	61
5.3.10.3 X-ray data collection and processing	61
5.3.10.4 Data analysis, model building, refinement, and structural analysis	61
6. Results	62
6.1 Studies on ERp28	62
6.1.1 Gene structure of mouse ERp28	62
6.1.2 The amino acid sequence of mouse ERp28 and the similarity with the human protein	63
6.1.3 RT-PCR and the possible alternative splicing form of mouse ERp28	64
6.1.4 Mouse genomic DNA library screening	67
6.1.5 Map of restriction sites in mouse ERp28 locus	67
6.1.6 Mouse ERp28 knock out vector composition	68
6.1.7 Steps for making the ERp28 Knock out vector	69
6.1.8 Selection of transfected ES cells by G418 and Gancyclovir and screening of ES cells by PCR	71
6.1.9 Genotyping of chimeric mice by PCR	71
6.1.10 Genotyping of +/- heterozygotes (F1) by PCR and Southern blot	72
6.1.11. Genotyping of -/- homozygotes (F2) by PCR and Southern blot	73
6.1.12. Immunostaining on paraffin sections of E12 embryos	77
6.2 Structural and functional analysis of Wind	79
6.2.1 SDS-PAGE gel and native gel electrophoresis of purified recombinant proteins His-Wind and Wind-His	79
6.2.2 Cross-linking of His-Wind and Wind-His with glutaraldehyde	79
6.2.3 Insulin reduction assay	80
6.2.4 Crystal structure of His-Wind	81
6.2.4.1 Overall structure	81
6.2.4.2 Dimer structure	82
6.2.4.3 The conserved residues on the protein surface show a distinct pattern	82
6.2.4.4 The electrostatic potential and the hydrophobic patches on the surface	84

6.2.4.5 Other features of the structure	85
6.2.4.6 Comparison of Wind and other PDI-related proteins	85
7. Discussion	88
7.1 Studies on ERp28	88
7.1.1 The gene of ERp28 contains three exons	88
7.1.2 ERp28 may have an alternative splice form among mammalian cells	88
7.1.3 ERp28 is expressed at high level in some tissues at stage E12 of mouse embryo development	89
7.1.4 The phenotype of ERp28 knock-out mice, a mystery to be resolved	91
7.2 Studies on Wind	92
7.2.1 Wind has two distinct domains and forms a homodimer	92
7.2.2 Is the dimerization functional?	94
7.2.3 Both the b and D domain are necessary for the localization of Pipe	95
7.2.4 The –CTGC- motif is redox inactive and not involved in the localization of Pipe	96
7.2.5 The possible substrate binding site on the surface of the b domain	97
7.2.6 The structure and the possible function of the D domain	99
7.2.7 Comparison of the b domain of Wind with the a- and b-domains of PDI	100
7.2.8 Wind provides very useful information for the study of ERp28	101

1. Summary

My thesis work includes two parts: firstly, generation of a knock out mouse of a gene, ERp28, which encodes a putative endoplasmic reticulum chaperone protein and secondly, the structural and functional analysis of Wind, the *Drosophila* homologue of ERp28.

To generate the ERp28 knock out mice, the gene structure of mouse ERp28 (exon distribution in the locus) was first established. Then a mouse genomic DNA library was screened using [³²P] labelled cDNA probe which was prepared by RT-PCR and two positive clones containing the mouse ERp28 genomic locus were obtained. The restriction map of the gene locus was determined. The ERp28 gene has a simple structure, with only three exons. The first exon contains the initiator ATG (the translation starting codon) and the third exon contains the terminator TGA (the stop codon). The third exon encodes almost the complete D domain, which is unique to the PDI-D subfamily. Interestingly, during the preparation of cDNA probe for screening by RT-PCR, I found a possible alternative splicing form of ERp28 in which the second exon was spliced out. This form could also be found in humans and rats, but not in *Drosophila*, indicating that it exists in a limited number of species. However, further studies have to be performed to confirm and characterize this alternative splice form. Based on the gene structure and the restriction enzyme map of the locus, a mouse ERp28 knock out (KO) vector was made and was transfected into ES cells. 3 positive recombination ES clones were obtained by selection with antibiotics and PCR-based screening. Two of these positive clones were introduced into mouse blastocysts by microinjection and the blastocysts were used to generate chimeric mice. Male chimeric mice were mated with female wild type mice (C57BL6N) to produce heterozygotes. The heterozygotes were then recrossed to generate -/- homozygotes. The chimeric mice, heterozygotes and homozygotes were genotyped and confirmed by PCR and/or Southern blot. So far no obvious phenotype of the homozygotes (-/-) have been found. Therefore, some detailed analyses will be carried out to clarify the function of ERp28. In addition, although ERp28 expresses in a variety of tissues and organs, revealed by immunohistological studies, it shows relatively higher expression levels in some specific tissues or cells, for example glial cells in the brain stem and choroid plexus of the further ventricle and the heart, at least in stage E12 embryos. This provides very useful information for the phenotypic analysis.

The second part of my work was the structure determination, by crystallography, of a PDI-D β protein. As earlier attempts to crystallize human ERp28 were unsuccessful (Dr. D.M Ferrari, personal communication), I concentrated my efforts on the *Drosophila* homologue,

Wind. Wind is essential for dorsal-ventral patterning in *Drosophila* embryos, and it has recently been shown to be crucial for the correct localization of Pipe, a key factor in the process of dorsal-ventral patterning.

The crystallization project, performed in collaboration with Dr. Qijun Ma, Prof. Dr. Isabel Uson and Prof. George M. Sheldrick at the Department of Structural Chemistry, University of Goettingen was successful. The 1.9 Angstrom crystal structure shows that Wind forms a homodimer. Each of the monomers contains two distinct domains, a N-terminal thioredoxin-like domain and a C-terminal D-domain. The dimer interface is contributed by the thioredoxin-like domain only, without participation of the D-domain. The N-terminal domain has a very characteristic thioredoxin fold, with $\beta\alpha\beta\alpha\beta\alpha\beta\alpha$ structural elements. The C-terminal D-domain forms a unique domain containing only 5 α -helices. Homodimerization yields a deep dimer cleft, which is negatively charged and large enough to hold a small peptide. The conserved residues form very distinct clusters on the top surface of Wind, indicating their importance for Wind function. Studies combining data from structural analysis, biochemical experiments and mutagenesis work suggest a candidate substrate binding site around Y55 in the surface of thioredoxin domain. We also show through insulin reduction assays and mutagenesis that although Wind carries a CTGC motif at its very N-terminus, this motif is redox inactive and is not necessary for the transporting of Pipe from the ER to the Golgi. These results indicate that Wind may be a redox-independent chaperone/escort protein. Furthermore, we have recently (Ma *et al.*, 2003) shown that although both the b- and D-domains are indispensable for the function of Wind, the D-domain of Wind can be replaced with that of mouse ERp28 without defect of the translocation of Pipe. This indicates a similar/same function of the D-domain of both proteins. The results from Wind have provided very useful information for clarification of the function of mammalian ERp28. Furthermore, the structure of Wind is the first crystal structure of PDI-related proteins in the ER. The structure and possible substrate-binding site revealed in this work may provide important clues on substrate binding sites and mechanisms of function of other PDI proteins.

2. Introduction

2.1 Endoplasmic reticulum (ER) and protein folding

In eukaryotic cells, proteins destined for the secretory pathway or the extracellular space enter the ER and are then transferred to the Golgi apparatus, where they are sorted according to their final destination. Two important changes occur to a protein in the ER; it is properly folded into its native conformation and post-translationally modified. A protein is translocated into the ER in unfolded form. Folding occurs as the protein enters the lumen and this process is associated with modifications such as N-linked glycosylation, disulfide bond formation, signal peptide cleavage or addition of a glycosphosphatidyl- inositide (GPI) anchor. These modifications are important for correct protein folding and/or protein function. For multimeric proteins, usually, proper oligomerization also takes place in the ER. Several exquisite mechanisms are provided by the ER to promote the unfolded nascent proteins to be properly folded, prevent them from aggregation and guarantee that only those proteins which are correctly folded can leave the ER and that those failing to fold or assemble will not proceed through the secretory pathway. First, the ER provides a suitable environment for the process, with optimal pH and redox conditions and a vast array of chaperones and foldases available, including members of protein disulfide isomerase family (PDI), BiP (immunoglobulin heavy chain-Binding Protein) (Kassenbrock *et al.*, 1988), GRp94 (Bose *et al.*, 1996), Calreticulin (Hebert *et al.*, 1996), Calnexin (Bergeron *et al.*, 1994), Peptidyl-prolyl isomerase (PPIase) (Freskgard *et al.*, 1992). Second, under situations where unfolded or misfolded proteins accumulate in the ER, a signalling pathway called the Unfolded protein responses (UPR) is activated to upregulate the expression of chaperones in the ER and to generate more chaperones to fulfil the requirement (Chapman *et al.*, 1998. Harding *et al.*, 2002). Third, the ER has a very stringent quality control system by which the folding state of proteins are monitored (Chevet *et al.*, 2001; Ellgaard and Helenius, 2003). Proteins with a correct conformation can leave the ER for the next station, in contrast, unfolded, incompletely folded or misfolded proteins will be retained and bound by molecular chaperones or foldases in the ER and undergo further attempts to achieve the native conformation. However, if these attempts fail, the proteins will enter the degradation pathway called ERAD (ER associated degradation) through which the unfolded or misfolded proteins will be retro-translocated from the ER to the cytosol and eventually degraded in the proteasome (Ellgaard and Helenius, 2001; Hampton, 2002; Kostova and Wolf, 2003). In all of these mechanisms, the chaperones and foldases occupy the critical positions. This introduction will focus on one of them, protein disulfide isomerase (PDI) and other members of the PDI family.

2.2 Protein disulfide Isomerase

Protein disulfide isomerase is a member of the thioredoxin superfamily and is highly abundant in the ER-lumen of a variety of tissues and organs and is highly conserved between species. It was first isolated from liver and characterized by the pioneer of protein folding, Christian Anfinsen in 1963 based on its ability to catalyse the refolding of ribonuclease A, an enzyme with 4 disulfide bridges (Golberger., *et al*, 1963). Since then, with the efforts of many scientists, several important functions of protein disulfide isomerase including its ability to catalyse the formation, reduction and isomerization of disulfide bonds (redox/isomerase function), chaperone activity, forming the subunit of the enzyme proly-4-hydroxylase (P4H) (Pihlajaniemi *et al.*, 1987) and microsomal triacylglycerol transfer protein (MTP) (Wetterau *et al.*, 1990) have been detected. Meanwhile, numerous proteins having functional and structural similarities to PDI have been discovered and recently all of these proteins are assigned to a new protein family - the protein disulfide isomerase family. These proteins have some common features. First, they all contain a N-terminal signal sequence for translocation to the ER; second, all of them reside in the ER and have a C-terminal ER retrieval or retention signal KDEL, or similar motif KEEL, HDEL, QDEL, etc, with the exception of calsequestrin which is a SR (sarcoplasmic reticulum) protein (MacLennan and Wong, 1971); third, all of these proteins contain at least one or more redox active or inactive thioredoxin domain with some sequence similarity to the cytosolic protein, thioredoxin; more importantly these domains are highly conserved at the three dimensional level, all of them exhibiting the same characteristic thioredoxin fold. As for protein disulfide isomerase, it is about 510 amino acids in length and is composed of four thioredoxin-like domains, a C-terminal retrieval signal KDEL and an acidic region in between. Two redox active domains a and a' with the active motif –CXXC– are separated by two redox inactive domains b and b' without this motif (fig 2.1) (Ferrari and Söling, 1999; Freedman *et al*, 2002).

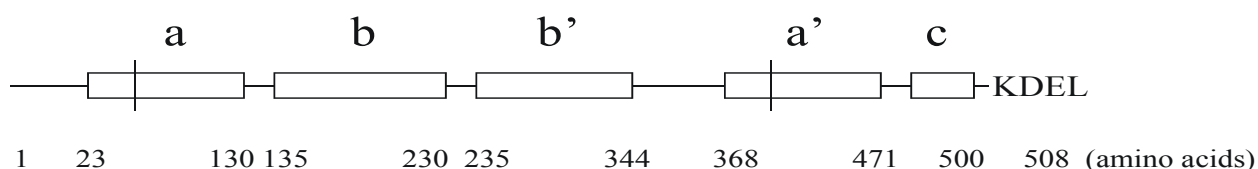


Figure 2.1: PDI model. Boxes represent a, b, b', a' domains and C-terminal acidic region, respectively. Vertical lines are referred to the CGHC redox-active motif. (see details in text).

2.2.1 Redox / Isomerase function of PDI and the pathway of formation of disulfide bonds

The formation of biosynthetic disulfide bonds is an important step in the maturation of both membrane and secreted proteins in eukaryotic and prokaryotic cells. Disulfide bridges are often vital for the stability of a final protein structure, the incorrect pairing of cysteine residues usually prevents the folding of a protein into its native conformation. Recent evidence indicates that disulfide bonds are not only an inert structural motif to enhance the stability of proteins but also a regulator for the activity of the mature protein by a thiol-disulfide bond exchange reaction (Hogg, 2003). For example, the cytoplasmic enzyme ribonucleotide reductase becomes oxidized during its catalytic cycles, and it must be recycled to its reduced form to be reactivated (Jordan and Reichard, 1998). In plants, light-generated reducing equivalents are used to reduce the regulatory disulfide bonds in several photosynthetic enzymes, thereby inducing a switch from catabolic to anabolic respiration (Dai *et al.*, 2001). Under specific cellular conditions, several transcription factors, including the bacterial OxyR (Helmann, 2002) and Hsp33 (Jakob *et al.*, 1999), also become activated by the oxidation of cysteines that form disulfide bonds (Fränd, 2000; Sevier and Kaiser, 2003). Thrombospondin-1 (TSP-1), a secreted extracellular glycoprotein participating in cell-cell and cell-matrix communication, has different disulfide-bonded forms *in vivo* and the disulfide interchange may be important for controlling its function (Essex and Li, 1999; Essex, 2001; Hogg, 2003).

The formation of a disulfide bond *in vitro* is a simple process during which two cysteines form a disulfide bond by transferring their electrons to an available acceptor, such as molecular oxygen. *In vivo*, however, the major mechanism for the formation of protein disulfide bonds is a thiol-disulfide exchange reaction of free thiols (-SH) with an already disulfide-bonded species. A thiol-disulfide exchange reaction can occur between a protein and any sulfhydryl-containing substrate, including small thiol-containing compounds, such as glutathione, or a protein containing a disulfide bond. This process is catalysed by a class of proteins commonly known as thiol-disulfide oxidoreductases including DsbA (Bardwell *et al.*, 1991), DsbC (Missiakas *et al.*, 1993; Bardwell *et al.*, 1993) in *E.coli* and protein disulfide isomerase and PDI-like proteins in eukaryotes. The activity of these proteins depends on a pair of cysteines that are often arranged in a -CXXC- motif within their thioredoxin domain (Sevier and Kaiser, 2003; Hiniker and Bardwell, 2003).

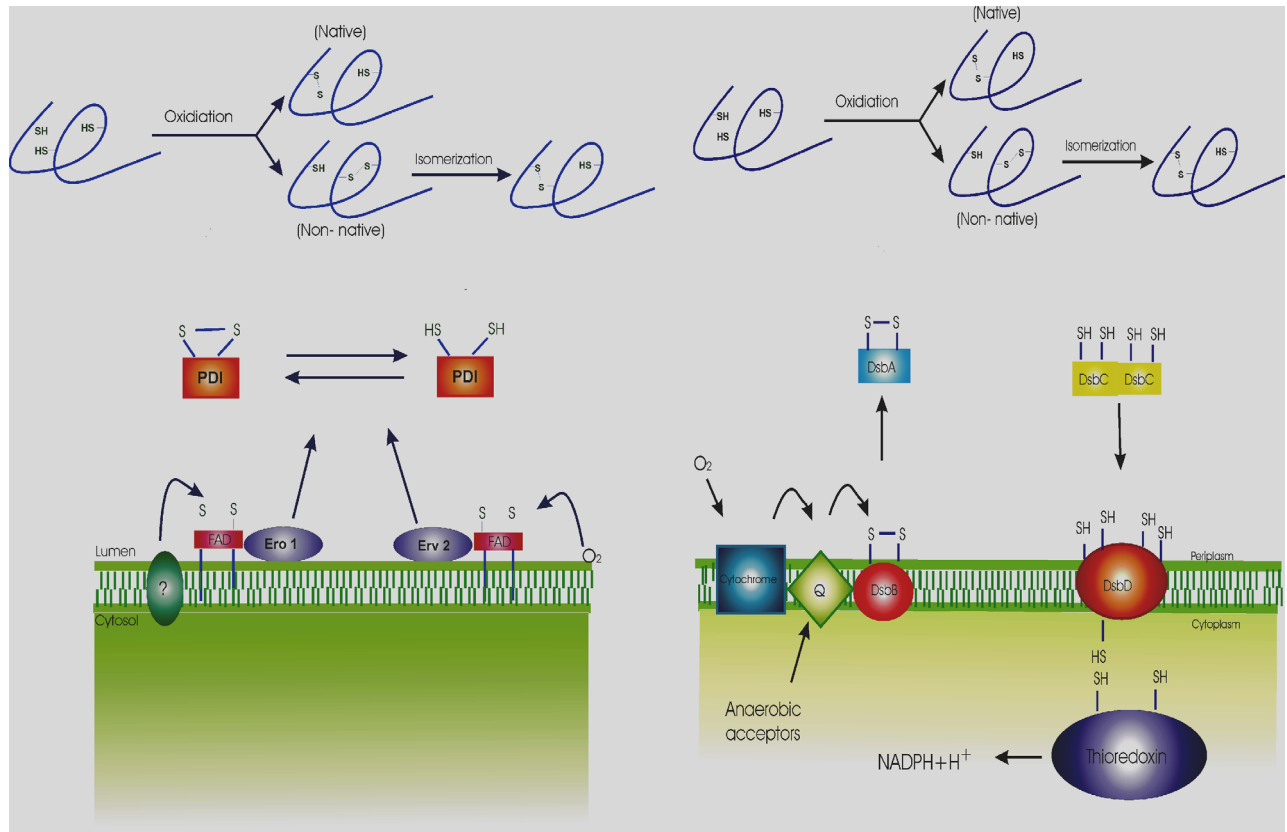


Figure 2.2: The pathway of disulfide bond formation in eukaryotes and prokaryotes (see details in text).

In the *E.coli* periplasm, disulfide bond formation and disulfide bond isomerization are catalysed by two separate pathways. The DsbA-DsbB pathway oxidizes thiol groups to form disulfides, while the DsbC-DsbD pathway isomerises mismatched disulfides. In the oxidative pathway, DsbA interacts with a folding protein containing reduced cysteines, oxidizing them to form disulfide bonds. In this process, DsbA is reduced and must donate its electrons to the inner membrane protein DsbB in order to be reoxidized. Under aerobic conditions, DsbB donates electrons to ubiquinone, which passes them to cytochrome oxidases and finally to molecular oxygen (Bader *et al.*, 1998, 1999, 2000). Under anaerobic conditions, DsbB donates its electrons to menaquinone, which donates them to anaerobic electron acceptors such as fumarate reductase or nitrate reductase (Bader *et al.*, 1998, 1999, 2000). In the DsbC-DsbD pathway, DsbC interacts with substrate proteins that contain non-native disulfide bonds and reshuffles these disulfide bonds to their native pairing (Zapun *et al.*, 1995; Rietsch *et al.*, 1996). DsbC requires DsbD in order to remain reduced in the highly oxidizing environment of the periplasm. The ultimate source of DsbD's reducing potential is cytoplasmic NADPH, which transfers electrons to thioredoxin, the cytoplasmic protein that directly reduces DsbD (Katzen and Beckwith, 2000; Sevier and Kaiser, 2003; Hiniker and Bardwell, 2003).

In eukaryotes, although it has been known for a long time that PDI and PDI-like proteins can catalyse the formation and isomerisation of disulfide bonds of substrate proteins, it was not clear how they could be re-oxidized until the discovery of Ero1 (Frand and Kaiser, 1998; 1999). Ero1 is a glycosylated luminal ER protein that is tightly associated with the ER membrane. It delivers oxidizing equivalents directly to PDI by consuming oxygen in a FAD (flavin adenine dinucleotide)–dependent reaction, which, in turn oxidizes the substrate proteins (Tu and Weissman, 2000; Tu and Weissman, 2002. Bardwell, 2002). Recently, in *Saccharomyces cerevisiae*, Erv2, a new membrane associated ER protein with a luminal domain that is non-covalently bound to FAD was discovered. It was found that Erv2 could form a mixed-disulfide intermediate with PDI and it seems that Erv2 could also drive the oxidation of substrate proteins by transferring oxidizing equivalents to PDI (Gerber, 2001; Sevier, 2001). However, whether mammalian homologues of Erv2 exist is not clear. These findings and the observation that yeast mutants deficient in glutathione synthesis owing to mutations in the GSH1 gene have little or no defects in protein disulfide bond formation have led to a revised view that protein-oxidation pathways proceed by the direct transfer of oxidizing equivalents between enzymes and do not rely on oxidizing equivalents provided by glutathione, a long term candidate (Cuozzo and Kaiser, 1999). By the processes above, PDI can maintain its oxidized form which is dominant in the ER, although the reduced form is observed as well (Tu *et al.*, 2000; Mezghrani *et al.*, 2001). Interestingly, protein disulfide isomerase alone can act both as an oxidase to form disulfide bridges in newly synthesized proteins, as well as an isomerase to reshuffle wrongly formed disulfide bridges. In contrast, in *E.coli*, the two tasks are assigned to two proteins, namely DsbA and DsbC, respectively.

The redox activity of PDI is dependent on the redox potential of the medium, reflected in the ratio of reduced to oxidised glutathione in the ER (GSH:GSSG ca 1:1-3:1, whereas 30:1 in the cytosol) (Hwang *et al.*, 1992). The standard redox potential of PDI is about –180 mV, much more oxidizing than that of thioredoxin (-260 mV) (which functions as a protein disulfide reductase in the cytosol) but less than that of bacterial DsbA (-100 mV) (Hawkins *et al.*, 1991). The difference in redox potential is due mainly to the nature of the two intervening residues of the reactive –CXXC– sequences (Huber-Wunderlich and Glockshuber, 1998). In consistence with the redox potential, PDI is about 50-fold more active than thioredoxin at catalyzing the isomerization of disulfide bonds in scrambled RNase (Hawkins *et al.*, 1991). Because it is clear now that oxidized glutathione is not required to re-oxidize PDI, it is speculated that glutathione functions as a net reductant in the ER to buffer the oxidizing

activity of Ero1. It is also possible that glutathione reduces the normally oxidized PDI, thus shifting PDI activity from oxidizing to isomerizing activity (Fassio and Sitia, 2002).

2.2.2 Chaperone function:

Molecular chaperones were originally defined as proteins that assist in the self-assembly of other polypeptide chains but are generally not part of the final functional units (Ellis and van der Vies, 1989). Thus, by definition, classical molecular chaperones interact only transiently with their substrate proteins. Within the ER, the term molecular chaperone has generally been used to describe chaperones that aid other proteins in folding and/or that retain unassembled proteins. However, the definition of an ER molecular chaperone is expanding and now includes roles for such proteins in the transport of substrate to a particular subcellular compartment as well as the modulation of substrate activity (Hendrick and Hartl, 1993; Bu and Schwartz, 1998). Roughly, there are four different types of proteins which adapt to this expanded definition. (1) classical ER chaperones like BiP, GRP94, Calreticulin, Calnexin; among them Calreticulin and Calnexin are primarily involved in the folding of glycoproteins; (2) Proteins such as PDI function as both classical ER chaperones and foldases; (3) Protein-specific ER chaperones such as RAP, MTP; (4) molecular escort proteins such as RAP which routinely accompany their substrate proteins out of the ER (Kim and Arvan , 1998; Bu and Schwartz, 1998).

The foldase activity of protein disulfide isomerase has been known well since it was shown to catalyse the formation and reshuffling of the disulfide bonds in substrates (Freedman *et al.*, 1994). The chaperone activity of PDI is proven by its capacity to promote the folding of denatured D-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and rhodanese (Cai *et al.*, 1994; Song and Wang, 1995), which have no disulfide bonds. This shows that the chaperone activity of PDI can be independent of the redox/isomerase activity, although chaperone activity has also been shown with disulfide-containing proteins such as lysozyme (Puig and Gilbert, 1994). Indeed, the chaperone activity may be necessary for PDI to fulfil its redox/isomerase function, because in order to promote the joining of thiol groups distantly situated in the peptide sequence to form correct disulfides, the peptide chain has to be folded at least to some extent to bring the thiol groups together. However, it remains unclear to what extent the chaperone activity of PDI is necessary for the redox/isomerase activity. Recently, PDI has been described to function as a redox-driven chaperone in the unfolding of the A1 chain of cholera toxin in the ER lumen, prior to its transport to the cytoplasm (Tsai *et al.*, 2001). In this hypothesis, PDI binds and unfolds the substrate in its

reduced state. The complex is then targeted to the ER membrane, where it binds to a protein at the luminal side of the membrane. Oxidation of PDI by Ero1p releases the substrate, possibly directly into the retrotranslocon (Tsai *et al.*, 2001; Tsai and Rapoport, 2002; Tsai *et al.*, 2002). However, this redox-dependent mechanism has been questioned by Lumb and Bulleid (2002) who observed that PDI binding and release to two other substrates was not driven by the redox state of the protein.

2.2.3 Other functions:

2.2.3.1 Subunit association

The enzyme prolyl-4-hydroxylase (P4H), an ER-luminal soluble protein that catalyses procollagen pro- α chain prolyl hydroxylation, is a heterotetrameric ($\alpha_2\beta_2$) enzyme (Pihlajaniemi *et al.*, 1987). Microsomal triacylglycerol transfer protein (MTP) is heterodimeric ($\alpha\beta$) and facilitates incorporation of triglycerides into lipoproteins (Wetterau *et al.*, 1990). In both proteins, PDI acts as the β subunits and probably functions as a chaperone to retain or stabilize the complex in the ER, as irreversible inactivation and aggregation ensues upon removal of the β -subunit. However, the chaperone activity is not its only function, as coexpression of the P4H catalytic α -subunit with BiP can generate a soluble α -subunit-BiP complex, which, however, has no P4H activity (Veijola *et al.*, 1996). In addition, for both enzymes, the redox / isomerase activity of PDI is dispensable since mutagenesis of both redox active sites of PDI subunit has no effects on activity (Lamberg *et al.*, 1996).

2.2.3.2 Different Localization of PDI

Despite the integrity of the KDEL signal, some reports show additional localizations of PDI including the Golgi, secretory vesicles, plasma membrane, and even cytosol in some cell types (Akagi *et al.*, 1988). In glial astrocytes, a significant amount (25%) of the total cellular PDI is claimed to be present in the cytosol (Safran and Leonard, 1991). Upon exposure of glial cells to thyroxine, the cytosolic PDI redistributes almost completely to become actin-associated, suggesting that PDI may acts as a mediator of thyroxine-induced actin nucleation (Farwell *et al.*, 1990; Safran *et al.*, 1992). Secretory cells such as hepatocytes and pancreatic acinar cells can saturate the –KDEL retrieval mechanism, resulting in the non-covalent binding of secreted PDI to the plasma membrane (Terada *et al.*, 1995). The appearance of the secreted PDI at the cell surface may have significant physiological effects. First, localization of PDI to the plasma membrane may elicit an autoimmune response. The

Long Evans Cinnamon rat displays a spontaneous hereditary hepatitis and hepatic carcinoma. The autoimmune antibodies generated in the progression of this hereditary hepatitis bind PDI and calreticulin (Yokoi *et al.*, 1993). Anti-PDI antibodies are also found in humans with liver disease and chronic alcoholism. Second, membrane-associated PDI may participate in the reduction of membrane receptors and / or ligands (Kroning *et al.*, 1994). CD4 is a member of the immunoglobulin (Ig) superfamily of receptors that mediates cell-cell interactions in the immune system and is the primary receptor for HIV-1. HIV-1 binds to CD4 via its gp120 envelope protein. This binding leads to interaction of the complex with a chemokine receptor, triggering fusion of the viral and cell membranes, leading to HIV-1 entry and infection (Eckert and Kim, 2001). Thioredoxin cleaves the disulfide bond in the second Ig-like domain of CD4 and PDI can cleave two of the nine disulfide bonds in gp120 after the binding of gp120 to CD4 cells. These cleavages might be important for conformational changes in CD4 and gp120 which are required for fusion of the viral and cell membranes (Matthias and Hogg, 2003). This is highlighted by the finding that mono- and dithiol alkylating agents which inactivate thioredoxin and PDI, and react with reduced CD4 and gp120- inhibit HIV-1 entry and envelope-mediated cell-cell fusion (Ryser *et al.*, 1994; Matthias *et al.*, 2002). Anti-PDI monoclonal antibodies also inhibit HIV-1 entry and cell-cell fusion (Gallina *et al.*, 2002). Cell surface PDI may also participate in cell signalling through interactions with membrane receptors as reported by Couet *et al* (1996) who showed that membrane PDI catalyses the shedding of the extracellular domain of the TSH receptor in the thyroid gland.

2.2.4 The structure of PDI and its functions

Since the first cDNA of PDI was sequenced, internal sequence homologies within the protein have been recognized and a multi-domain protein architecture proposed. The current model is based on the combination of data from proteolysis of native PDI and characterization of recombinant fragments, bioinformatic approaches, and NMR analysis. PDI comprises of four structural domains, a, b, b', a', plus a linker region between b' and a' and a C-terminal acidic region which is a putative low-affinity, high-capacity Ca^{2+} -binding site (fig 2.1) (Ferrari and Söling, 1999; Freedman *et al.*, 2002). Both a and a', with high sequence similarity to thioredoxin, contain a redox active site –CGHC-. b and b' domains, without the –CXXC- active site, show sequence similarity to each other but no obvious similarity to thioredoxin or to the a' domain. However, NMR studies on the recombinant a and b domains of human PDI have clearly revealed that both of them form a characteristic thioredoxin fold (Kemink *et al.*, 1995, 1996, 1999).

2.2.4.1 Structural comparison between thioredoxin and a and b domains of PDI

Thioredoxin is a small (12 kDa), ubiquitous protein with a disulfide reductase function in the more reductive environment of the cytosol. It contains a –CGPC– site which is crucial for the reduction of disulfide bonds (Holmgren, 1985, 1989). Most, but not all proteins catalysing the redox reactions involving reactive dithiols *in vivo* belong to the thioredoxin superfamily. The large family accommodates thioredoxin-like, glutaredoxin-like and PDI-like proteins, as well as members of the bacterial Dsb family (Gilbert, 1998; Holmgren, 1989; Bardwell, 1994). The single domain members of the thioredoxin superfamily (thioredoxins and glutaredoxins) all have the same α/β fold, with the structure $\beta\alpha\beta\alpha\beta\beta\alpha$. A five-stranded β -sheet, with all of the strands except β_4 being parallel, forms the central core, surrounded by four α -helices (Katti *et al.*, 1990; Eklund *et al.*, 1992). The –CXXC– redox active motif is found in an exposed turn linking β_2 to α_2 . The sulfur atom of the more N-terminal Cys residue is at the N-terminal pole of the α_2 and is exposed at the surface of the molecule, while the sulfur atom of the more C-terminal Cys residue is buried behind it. The core thioredoxin structure, $\beta\alpha\beta\alpha\beta\beta\alpha$, can also be detected in other members of the thioredoxin superfamily, including DsbA (a bacterial periplasmic oxidase) (Martin *et al.*, 1993), and in a number of enzymes involved in glutathione or sulfur metabolism (Martin, 1995). In thioredoxin, the first cysteine of the –CXXC– site, C32 has a pKa of 7.1, much lower than that of the free cysteine (pKa of 8.7 at neutral pH), rendering it highly reactive. At neutral pH the reactive sulfur atom of C32 may share a hydrogen bond to the –SH hydrogen

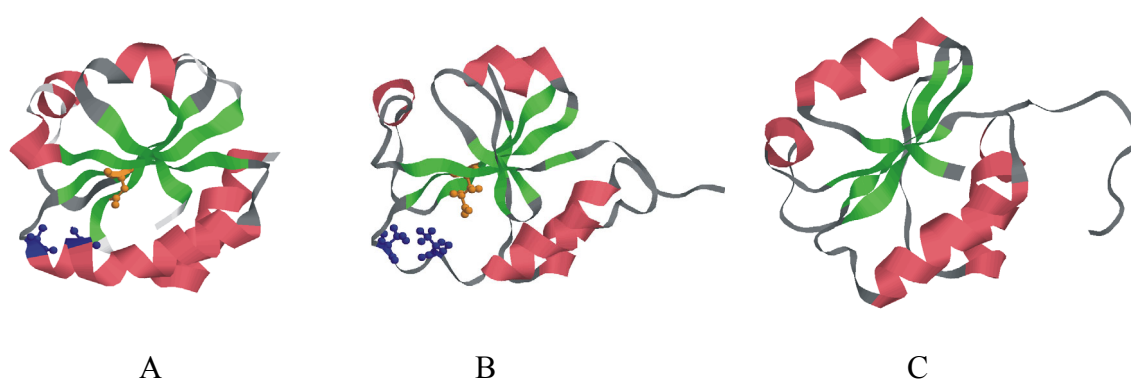


Figure 2.3: Comparison of the tertiary structure of thioredoxin with the thioredoxin fold of the a and b domains of PDI. (A) *E. coli* thioredoxin ribbon model of the crystal structure determined by Katti *et al.* (1990). (B) The a-domain of human PDI, based on a model suggested by NMR (Kemmink *et al.*, 1997). The Cys residues of the active site (blue), D26 (yellow) of thioredoxin (A) and E30 (yellow) of the a-domain of PDI (B) are shown as ball and stick representation. (C) The b-domain of human PDI, based on a model suggested by NMR (Kemmin *et al.*, 1999). α -helical elements are shown in red, and β -strands in green. (see details in text).

of C35. The pKa of C32 is thought to be decreased by a nearby buried partial charge on D26 (Chivers *et al.*, 1996), which may serve as a general acid/base in thioredoxin-catalysed redox reactions. The C32 thiolate can make a nucleophilic attack on disulfides, generating a mixed disulfide that is then disrupted by C35 to produce a reduced substrate protein. The reactivity of C35 follows upon its loss of a proton to D26 (fig 2.3 A). The thiolate species generated is well poised to attack the intermolecular disulfide-bonded C32. Thioredoxin is then recycled to the reduced state by thioredoxin reductase and NADPH. In close proximity to the reactive-site pocket are the hydrophobic residues G33, P34, I75, P76, G92 and A93, which may be important for protein-protein interaction (Holmgren, 1995).

The a domain of human PDI, in addition to adopting the small overall α/β fold, shares many common features with thioredoxin (fig 2.3 B); first, the active site motif is located at the N-terminus of helix α_2 , which is distorted by a proline residue; second, the peptide bond before the proline residue at the N-terminus of β_4 is in the *cis* conformation; third, there is a buried acidic residue, E30, in analogous position in the a domain of PDI to that of D26 in thioredoxin (Kemink *et al.*, 1995; 1996) (fig 2.3 B). As in the a-domain, the b-domain of PDI also forms a thioredoxin fold (Kemink *et al.*, 1999), but the characteristic thioredoxin-like active site has been deleted and other residues associated with redox properties have been replaced (fig 2.3 C). Despite a lack of 3-D structural data, based on the sequence similarity between a and a', b and b', and the secondary structure prediction and alignment, the a' and b' domains of PDI should also adopt the thioredoxin fold (Freedman *et al.*, 2002).

2.2.4.2 The importance of the b' domain of PDI

The redox / isomerase activities of PDI, as in thioredoxin, are due to the reactivity of the -CXXC- redox active sites in the a and a' domains. Redox assays using simple peptide substrates show that the isolated a and a' domains retain nearly full redox activity (Darby and Creighton, 1995a,b), however, neither of them could fulfill the isomerase activity of the full length PDI, as neither is able to catalyse the isomerisation of the disulfide bonds in BPTI (the bovine pancreatic trypsin inhibitor), lysozyme and ribonuclease (Darby and Creighton, 1995c). This suggests that other parts of PDI are required for its full range of activities. Recently, the b' domain of PDI attracted special attention. Through the intensive studies of a series of combinations of different domains, it was revealed that the simple thiol-disulfide reaction only requires the a or a' domains, that simple isomerization requires one of these in a linear combination of domains including b', while complex isomerization (isomerizations that would require substantive conformational change in the substrate as well as thiol-disulfide

reactions) requires all four of these PDI domains (but not the C-terminal acidic extension) (Darby and Creighton, 1995c; Darby *et al.*, 1998). The b' domain is also crucial for the peptide binding activity of PDI. For small peptides, the isolated b' domain is essential and sufficient for the binding. However, the binding of larger peptide or non-native protein substrates requires at least the b'-a'-c fragment. Thus it appears that the b' domain of PDI provides the principal peptide binding site of PDI, but that all domains contribute to the binding of larger substrates such as non-native proteins (Klappa *et al.*, 1998). These observations explain the apparent importance of the b' domain in the catalysis of complex isomerization reactions by PDI, suggesting that it holds the substrate protein in a partially unfolded conformation while the catalytic sites act synergetically to perform the chemical processes of thiol-disulfide exchange (Freedman *et al.*, 2002).

2.3 PDI family or PDI-L (like) family

In the last few years, the PDI family has expanded by the addition of several new members. Currently, 12 known mammalian proteins are likely to belong to this family. The number is much bigger if the homologues from other species are included. Most of these proteins are redox active and involved in thiol-disulfide exchanges with the notable exception of the most recently found members. Although no direct structural information is available for all of them, by sequence alignment and secondary structure prediction,

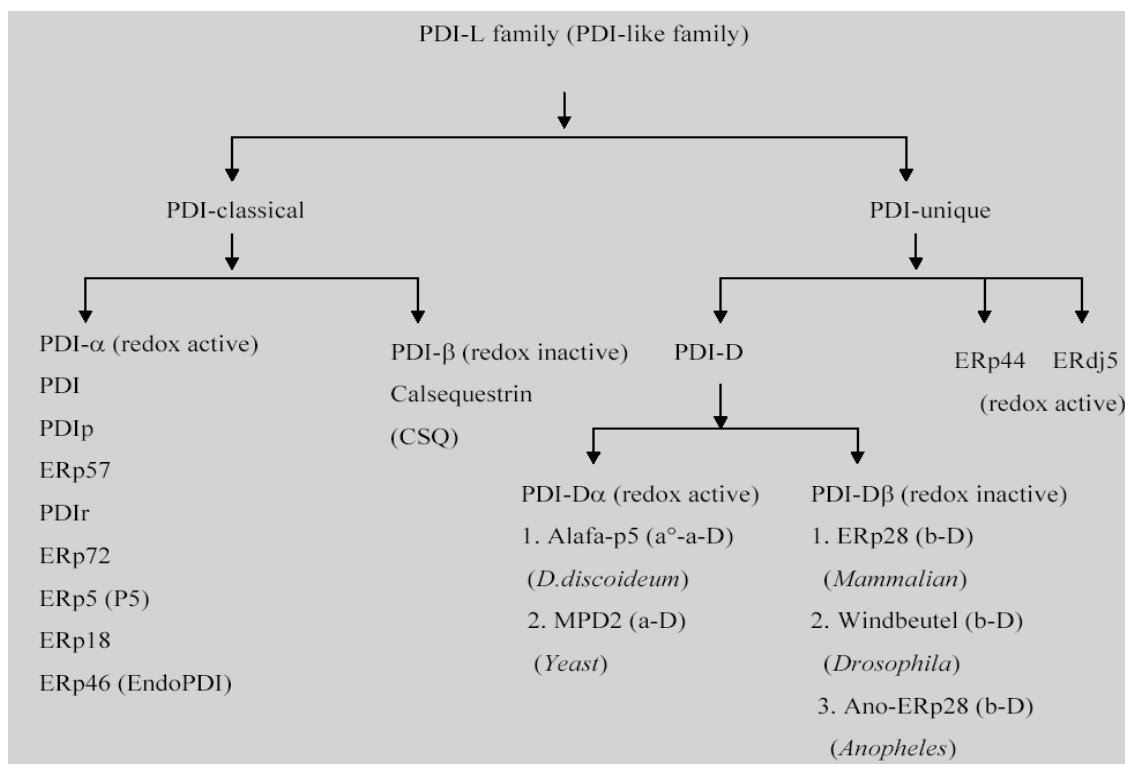


Figure 2.4: The PDI-L family.

Mammalian PDI-like proteins						
proteins	Size (kDa)	Domain composition	Acidic region	Active-site sequence	ER-localization signal	Unique features
PDI	55	a-b-b'-a'-c	1	CGHC	KDEL	Mixed disulfide with Ero1 General peptide-binding site
ERp57	54	a-b-b'-a'	-	CGHC	QDEL	Interaction with amino-glycosylated proteins through calnexin/calreticulin
PDIp	55	a-b-b'-a'	-	CGHC, CTHC	KEEL	Pancreas-specific expression
ERp72 (CaBP2)	71	c-a°-a-b-b'-a'	1	CGHC	KEEL	?
PDIr	57	b-a°-a-a'	-	CSMC, CGHC, CPHC	KEEL	?
P5 (CaBP1)	46	a°-a-b-c	1	CGHC	KDEL	Involved in left/right asymmetric patterning
ERp18	18	a	1	CGAC	EDEL/QDEL	?
ERdj5 (JPDI)	91	J-a-b-b'-a'-a''-a°	-	CSHC, CPPC	KDEL	Interaction with Bip
ERp44	44	a-CSQ-U	-	CRFS	RIEL	Mixed disulfide with Ero1
CSQ	53	b°-b-b'	-	-	-	Ca ²⁺ binding protein in SR
ERp28	28	b-D	-	-	KEEL	?
ERp46 (EndoPDI)	46	a°-a-a'-c	1	CGHC	KDEL	?

Table 2.1: Mammalian PDI-like proteins. a, a', a'' and a° represent redox-active thioredoxin domains; b, b' and b° are referred to redox-inactive thioredoxin domains; c is the acidic region; J the DnaJ-like domain; CSQ is CSQ (calsequestrin)-like domain; U a unique region (domain) without similarity to other proteins; D the α -helix domain of the PDI-D subfamily.

their domain organization can be deduced. Based on the domain compositions, the PDI family can be divided into a classical PDI-L subfamily which contains only thioredoxin domains (redox active or inactive) and a unique PDI-L subfamily which contains other unique domains apart from the thioredoxin domains. These subfamilies can be divided further into α (redox active) and β (redox inactive) sub-subfamilies (reviewed by Ferrari and Söling, 1999).

2.3.1 ERp57

ERp57 has the same modular structure of active and inactive domains as PDI, but lacks the C-terminal acidic region. At the amino acid level, overall identity with PDI is 33%. Like PDI, ERp57 shows a ubiquitous distribution in a wide variety of cell types. However, ERp57 exhibits a low redox activity and cannot substitute for PDI as the β subunit of P4H. ERp57 interacts with nascent monoglycosylated glycoproteins, but not with non-glycosylated proteins, in a disulfide-independent manner (Oliver *et al.*, 1997). This is in contrast with PDI, which interacts with proteins independently of their glycosylation status (Van der Wal *et al.*, 1998). It is clear now that the specificity of this interaction with N-glycosylated proteins is not intrinsic to ERp57 as ERp57 itself lacks lectin-like properties, but instead is due to a long-term association between ERp57 and calnexin or calreticulin. These glycoprotein-specific chaperones of the ER are responsible for substrate recognition / binding in the multi-subunit chaperone / ERp57 complex (Zapun *et al.*, 1998; High *et al.*, 2000). Hence, it appears that ERp57 does not possess a general binding site for non-native proteins, but rather that it has a specialized binding site for specific partner proteins (calnexin and calreticulin).

2.3.2 PDIp

PDIp, the pancreas-specific PDI family member, is the only known family member to date to show a highly specific tissue expression, being exclusively expressed in the acinar cells of the pancreas (Desilva *et al.*, 1997). It is also the only member of the PDI family known to be glycosylated. It has the same domain organization as human PDI, but like ERp57, lacks the acidic region. It shares 40-45% amino acid identity with PDI. The specific tissue distribution suggests that PDIp may be involved in the folding of only a subset of secreted proteins (*e.g.* pancreatic zymogens). Using cross-linking and competitive binding studies, PDIp has been shown to bind misfolded proteins (scrambled RNase A) and peptides (including zymogen-derived peptides) *in vitro*, and the motif defining the specificity of binding of peptides and other small ligands by PDIp has also been identified (Ruddock *et al.*, 2000; Klappa *et al.*, 2001). PDIp binds specifically to peptides that include either a Y or a W residue (except where the Y or W is C-terminal or where there is an adjacent negatively charged residue). Interestingly, the peptide binding interaction could be inhibited by competition using stoichiometric concentrations of oestrogens, such as 17 β -oestradiol (Klappa *et al.*, 1998).

It should be emphasized that PDI, PDIp and ERp57 all share a similar organization, with a linear sequence of four thioredoxin-like domains in an a-b-b'-a' pattern. The strong

overall homology between these three proteins suggests that they share a similar function, but there are significant sequence differences in the b' domains, indicating that PDIp and ERp57 are isoforms of PDI with specialized substrate binding properties.

2.3.3 ERp72 (CaBP2)

ERp72 contains three active thioredoxin domains rather than two as PDI. Like PDI, it is an abundant, ubiquitous, stress-inducible protein with calcium-binding capacity (Van *et al.*, 1988, 1993). Although ERp72 possesses significant redox and disulfide isomerase activity and can complement PDI-deficient yeast (Gunther *et al.*, 1993), the information about its concrete functions in vivo is still elusive. ERp72 and BiP have been shown to coprecipitate with an overexpressed, mutated substrate protein lacking glycosylation sites (human chorionic gonadotropin β -subunit) in a process that probably involves disulfide-bond formation and which is independent of calnexin (Feng *et al.*, 1996). The protein interacts in vitro with denatured proteins in association with molecular chaperones, including PDI, Bip and GRP94 (Nigam *et al.*, 1994). Similar findings have been made in vivo for thyroglobulin and another secretory protein, thrombospondin (Kuznetsov *et al.*, 1996). According to these findings, ERp72 resembles PDI rather than ERp57.

2.3.4 PDIr

PDIR is preferentially expressed in cells actively secreting proteins and that the expression of PDIR is stress-inducible (Hayano and Kikuchi, 1995). Although little is known about the function of PDIr, its domain structure (b-a^o-a-a') is interesting. Each a-type domain has a different active-site sequence: -CSMC-, -C-GHC- and -C-P-H-Cys-. It is speculated that each a domain has a different rate of catalysis, which may be optimal for the particular target substrates (Ferrari and Söling, 1999). Very recently, a new member ERdj5 with a specific DnaJ binding domain was identified that also contains several -CXXC- with different residues between the two cysteines (Hosoda *et al.*, 2003; Cunnea. *et al.*, 2003). Further studies are required to establish the role of the different active-site sequence, and to determine the extent to which they remain redox activity.

2.3.5 ERdj5 (JPDI)

ERdj5, a newly characterized member of PDI family, contains domains resembling DnaJ, PDI and thioredoxin. ERdj5 is a ubiquitous protein and is particularly abundant in secretory cells like most members of PDI family. *In vitro* experiments demonstrated that

ERdj5 interacts via its DnaJ domain with BiP in an ATP-dependent manner. Using the standard thioredoxin activity assays of insulin reduction, recombinant ERdj5 did not show any activity. However, this does not exclude that ERdj5 can act as a reductase on other substrates. The expression level of ERdj5 is upregulated during ER stress, suggesting potential roles for ERdj5 in protein folding and translocation across the ER membrane. It is postulated that the DnaJ domain of ERdj5 could stimulate the ATPase activity of BiP, allowing BiP to dissociate from interacting polypeptide, thus leaving the cysteine residues of the polypeptide available for ERdj5 (Hosoda *et al.*, 2003; Cunnea. *et al.*, 2003).

2.3.6 ERp44

By co-immunoprecipitation and mass spectrometry, a novel UPR induced ER protein, ERp44, was identified. ERp44 can form mixed disulfides with both human Ero1 homologues, Ero1-L α and Ero1-L β , as well as with partially unfolded Ig subunits. It contains a thioredoxin-like domain followed by a fragment displaying weaker similarities with the second domain of calsequestrin and a long sequence with no obvious sequence homologies with other proteins. In the thioredoxin-like domain, the second cysteine in the canonical –CXXC– motif is replaced by a serine, yielding the sequence –CRFS–. The CRFS motif and the surrounding sequences are extremely conserved in all species, suggesting an important functional role for this region. MS analysis indicates that the cysteine of the CRFS motif is involved in the formation of mixed disulfides with Ero1-L α . It has been shown that Ero1-L α displays two dominant redox isoforms, OX1 and OX2, which possibly represent different structural conformers of Ero1-L α in its functional cycle. Overexpression of ERp44 alters the relative amounts of the two isoforms, shifting the equilibrium towards OX2. By favouring the accumulation of OX2, ERp44 may play an important role in controlling the function of human Ero1 and hence the redox state of the ER (Anelli *et al.*, 2002).

2.3.7 ERp18

The recently identified ERp18 (Alanen *et al.*, 2003) is the smallest member of the PDI family. It contains a single thioredoxin domain and the –CGAC– motif resembles the –CGPC– site of thioredoxin instead of –Cys-Gly-His-Cys of PDI. Due to this fact, it was assumed that ERp18 is likely to be a reductase in the ER rather than an oxidase. However, the reduced form of ERp18 is more stable than its oxidised form, similar to PDI. From this point, it seems to be an oxidase. It is worth noting that there is a very acidic region within its thioredoxin domain.

This feature is unique among the PDI family although PDI has an acidic region at the C-terminus outside the a' domain.

2.3.8 Calsequestrin (CSQ)

Calsequestrin (CSQ), an acidic protein, resides in the lumen of the junctional terminal cisternae of the sarcoplasmic reticulum at high concentration (MacLennan and Wong, 1971). CSQ is the major Ca^{2+} storage protein of muscle and is capable of binding and releasing large quantities of Ca^{2+} rapidly. Each molecule of CSQ binds about 40 to 50 Ca^{2+} ions with a binding constant of about 1 mM under physiological conditions (MacLennan *et al.*, 2002). CSQ is the first member of PDI family whose crystal structure was solved (Wang *et al.*, 1998). Although sequence similarity between CSQ and thioredoxin or PDI is low, the structure shows that CSQ is made up of three domains, each with a redox inactive thioredoxin fold. The core of each of the three thioredoxins domains is a 5-strand β -sheet surrounded by 4 α -helices which contain the bulk of the acidic residues. However, the individual thioredoxin domain is not the Ca^{2+} binding unit. Ca^{2+} is largely bound through cross bridges which form in clefts between each of the three domains. Thus the β -sheets in the thioredoxin-fold domains provide the platform on which a large number of negative charges can be brought together so that rapid Ca^{2+} binding and release are possible. Ca^{2+} ions are bound on the protein surface and between protein domains, rather than in a loop as in EF-hand proteins. The Ca^{2+} binding sites in CSQ contain numerous acidic residues that comprise over 1/3 of the total residues in the protein. In line with the sequence differences in each domain, the number of acidic residues involved in Ca^{2+} binding in each domain ranges from 13 to 32. The most acidic and most variable sequence in different CSQs is the C-terminus, with the C-terminus of skeletal muscle calsequestrin containing 14 contiguous acidic amino acids. This sequence, however, is disordered and is only folded into the Ca^{2+} binding domain when dimers form between two calsequestrin molecules, likely creating cross bridges that stabilize the dimer interface (Wang *et al.*, 1998; MacLennan *et al.*, 2002). As the major function of CSQ is Ca^{2+} binding, it is not surprising that it contains many acidic residues in the sequence which distinguishes it from other PDI members. However, the structural composition of the three thioredoxin domains still provides very helpful information for understanding the relationships between structure and function of other members.

2.3.9 P5 (CaBP1)

Using purified recombinant proteins, the redox/isomerase activity of human P5 and its rat homologue (CaBP1) have been confirmed, but the activity is lower than that of human PDI (Rupp *et al.*, 1994; Kikuchi *et al.*, 2002; Kramer *et al.*, 2001). Moreover, human P5 was observed to have peptide-binding ability, and its chaperone activity was revealed with rhodanese and citrate synthase as substrates, but not with D-glyceraldehyde-3-phosphate dehydrogenase, indicating that hP5 has substrate specificity with respect to chaperone activity (Kramer *et al.*, 2001; Kikuchi *et al.*, 2002). Mutation of two thioredoxin-related motifs in hP5 revealed that the first motif is more important than the second for isomerase activity and that the first cysteine in each motif is necessary for isomerase activity (Kramer *et al.*, 2001). Like PDI, the isomerase and chaperone activities of hP5 are probably independent, since mutation of the thioredoxin motif of hP5 retains its chaperone activity but lacks isomerase activity. Very interestingly, the zebrafish homologue of P5 was recently reported to be specifically involved in the process of establishing the left / right asymmetry in zebrafish embryos. The gene for zebrafish p5 is expressed at high levels in the organizer and axial mesoderm and is essential for establishing left/right asymmetries. Depletion of p5 production with morpholino oligonucleotides results in loss of the asymmetric development of the heart, liver, pancreas and gut. In addition, p5 depletion results in bilateral expression of all genes known to be expressed asymmetrically in the lateral plate mesoderm and the brain during embryogenesis (Hoshijima *et al.*, 2002). The *Drosophila* homologue DmCaBP1 is also expressed in a specific spatiotemporal pattern during embryogenesis. In particular, it is expressed in midline precursor cells in the developing CNS (Li *et al.*, 1998). These data indicate that p5, similar to *windbeutel* (see below), is very important during development, although it is not clear how p5 carries out its functions.

2.4 PDI-D subfamily

The classical members of the PDI family consist of thioredoxin-like domains, redox active or inactive. Recently, several members with domains different from thioredoxin-like domains have been identified including ERp28 (ERp29), ERp44 and ERdj5 (JPDI) in mammalian cells. ERp28 was discovered some years ago (Demmer *et al.*, 1997; Ferrari *et al.*, 1998). It contains not only a redox inactive thioredoxin-like domain at the N-terminus but also a novel domain, the D domain at the C-terminus. This D domain with about 100 amino acids in length is composed only of 5 α -helices (Ferrari *et al.*, 1998; Ferrari and Söling, 1999). Later on, homologues of ERp28 were found in many species. What is very interesting is that

in different species, the redox activity of these homologues are obviously different. For example, in mammals, ERp28 (ERp29) is definitely redox inactive because of the lack of a –CXXC– active site, in contrast, mpd2, the *S.cerevisiae* homologue, is redox active (Tachikawa *et al.*, 1997). Redox active homologues also exist in plants (Monnat *et al.*, 1997). These proteins are grouped into the PDI-D subfamily due to the presence of similar D domains at the C-termini in their sequences (Ferrari *et al.*, 1998; Ferrari and Söling., 1999). According to their redox activity, they are further divided into PDI-D α (redox active) and PDI-D β (redox inactive). Although the *Dictyostelium* homologue of ERp28 has no KEEL at its C-terminus, it can be efficiently retained in the ER (Monnat *et al.*, 2000). Furthermore, the last 57 C-terminal residues of Dd-PDI can sufficiently localize a green fluorescent protein (GFP) chimera to the ER (Monnat *et al.*, 2000). This indicates that the D domain may carry out a retention/retrieval function.

PDI-D subfamily

Proteins	Size	Domain structure	ER localization sequence
PDI-Dα			
<i>S. cerevisiae</i> mpd2	277	a-D	HDEL
<i>A. niger</i> Tig A	359	a°-a-D	KDEL
<i>N. crassa</i> ER38	369	a°-a-D	KEEL
<i>M. sativa</i> P5	364	a°-a-D	?
<i>N. tabacum</i> PDI	359	a°-a-D	?
<i>D. discoideum</i> Dd-PDI	363	a°-a-D	?
PDI-Dβ			
<i>H. sapiens</i> ERp28/29	216	b-D	KEEL
<i>R. norvegicus</i> ERp28/29	260	b-D	KEEL
<i>M. musculus</i> ERp28/29	260	b-D	KEEL
<i>D. melanogaster</i> Wind	257	b-D	KEEL
<i>Anopheles</i> Ano-ERp28	270	b-D	RAEL

Table 2.2: PDI-D subfamily. PDI-D α proteins are typically 360 amino acids in length and have two redox active thioredoxin domains followed by a D-domain except that mpd2 has only one. PDI-D β proteins, in contrast, are composed of a single, redox inactive thioredoxin domain followed by a D-domain. Sizes are given in amino acids residues. ? means that an ER retrieval signal is missing, so that ER-localization may solely depend on the D-domain.

2.4.1 ERp28

Human ERp28 (rat ERp29) is ubiquitously expressed in tissues and cells. It shows higher expression levels in secretory tissues and cells (Shnyder *et al.*, 2002). As mentioned above, it is composed of two distinct domains, the N-terminal thioredoxin domain and the C-terminal D-domain. As the thioredoxin domain lacks the –CXXC– redox active motif, apparently, ERp28/ERp29 does not have the redox/isomerase activity that PDI has. However, although ERp28 cannot assist the refolding of denatured GAPDH and denatured, reduced RNase, it does coprecipitate with overexpressed HBS (hepatitis B small antigen) (Ferrari *et al.*, 1998). The results suggest that ERp28 is not likely to be a general chaperone but that it may be a specific chaperone with a narrow substrate specificity or perhaps function as an escort protein like RAP (Bu *et al.*, 1995; 1998; 2001). This is supported by the recent finding that the rat homologue, ERp29 is associated with proteins forming the thyroglobulin folding complex in the ER of the thyroid epithelial cells (Sargsyan *et al.*, 2002). Folding and post-translational modification of the thyroid hormone precursor, thyroglobulin, is facilitated by several molecular chaperones and folding enzymes, such as BiP, GRP94, calnexin, protein disulfide isomerase, ERp72, and others. They have been shown to associate simultaneously and/or sequentially with thyroglobulin in the course of its maturation, thus forming large heterocomplexes in the ER of thyrocytes. In addition, during differentiation of mouse testicular carcinoma F9 cells, ERp28 protein levels rise at least 12-fold (Ferrari *et al.*, 1998). This upregulation during differentiation indicates that ERp28/ERp29 may be involved in the process of development similar to its *Drosophila* homologue Wind (Konsolaki and Schupbach, 1998).

2.4.2 Wind and the Dorsal-Ventral patterning of the *Drosophila* embryo

In *Drosophila*, the establishment of dorsoventral (DV) polarity occurs during oogenesis and requires communication between the germ-line derived oocyte and the somatically derived follicle cells of the ovary. The DV patterning process is launched by the communication from oocyte to the follicle cells, called the Gurken-EGFR pathway. Initially, the oocyte nucleus first moves to the anterior dorsal part of the cell and the gurken mRNA is synthesized between the oocyte and follicle cells. The product of gurken, Gurken (Neuman-Silberberg and Schupbach, 1993) which is homologous to epidermal growth factor (EGF), accumulate around the oocyte nucleus and is then secreted to the follicle cells, which differentiate to a dorsal morphology later. This signal is received by the follicle cells via

Torpedo, the homologue of the human epidermal growth factor receptor (EGFR) (Wadsworth *et al.*, 1985). Torpedo (EGFR) is expressed in all follicle cells, however, it is only activated in the dorsal follicle cells receiving the Gurken signal. This signal transduction leads to two major consequences: a change in the properties of follicle cells that prevents them from acquiring ventral fates and regulates the second pathway, especially by restricting the expression of the pipe gene (in the second pathway) in ventral follicle cells.

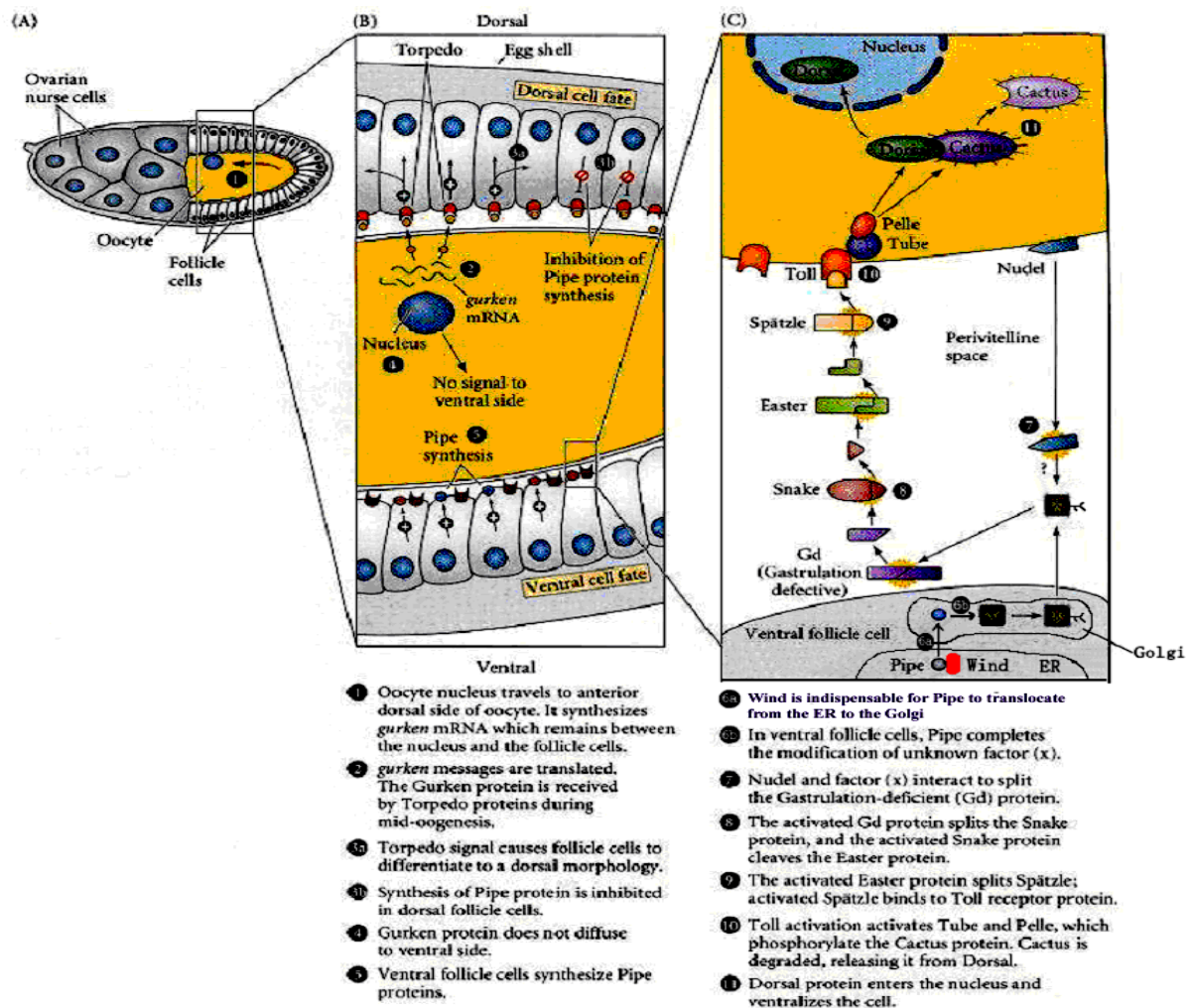


Figure 2.5: Dorsal-Ventral patterning of the *Drosophila* embryo (excerpted from Molecular biological course of Fritz Aberger and modified by Dr. Qingjun Ma) (See details in text).

The second pathway, which sends signals from the follicle cells to the embryo, requires the actions of a dorsal group including at least 11 genes and ultimately leads to the specification of the dorsal-ventral axis of the embryo. This process is realized via a proteolytic cascade (Morisato and Anderson, 1995), which results in the formation of a nuclear gradient

of the transcription factor Dorsal. Dorsal mRNA is supplied by the mother but protein gradient is generated after fertilization.

Windbeutel, *pipe* and *nudel* are the only three genes expressed by mother follicle cells. The gene, *windbeutel*, encodes a putative ER resident PDI-related protein, Wind, which is required to localize Pipe to the Golgi apparatus (Konsolaki and Schüpbach, 1998; Sen *et al.*, 2000). The *pipe* gene encodes the protein Pipe, the homologue of mammalian 2-O-sulfotransferase (2-OST) (Sergeev *et al.*, 2001). Pipe is the critical factor to define the DV symmetry (Sen *et al.*, 1998). The gene *nudel* encodes a modular protein with an extracellular matrix domain and a serine protease domain (Hong and Hashimoto 1995). It has been suggested that Nudel is secreted by the follicle cells and may possibly be incorporated in the vitelline membrane, thus specifying the site of generation of the active Spätzle ligand, after fertilization of the oocyte. Wind and Nudel are not spatially restricted to the ventral follicle and expressed in all follicle cells. However, the expression of Pipe is negatively regulated by EGFR. Thus Pipe is spatially restricted to the ventral follicle cells and its expression in dorsal follicle cells is inhibited by EGFR which is only activated in the dorsal side (Sen *et al.*, 1998). In the Golgi apparatus, Pipe modifies an as yet unidentified proteoglycan x, which combines Nudel to trigger a protease cascade leading to the DV polarization of the embryo. Thus, Pipe plays a pivotal role in the process that defines the DV axis of the embryo and its spatially regulated activity may provide the link between the establishment of DV polarity in the follicle and the transmission of DV patterning information to the developing egg and future embryo (Sen, 1998).

Gastrulation-defective (Gd), Snake, Easter are expressed in the oocyte and are all serine proteases. They are released into the perivitelline space. Gd is cleaved by Nudel/x and activated. Subsequently, Gd cleaves Snake, and Snake cleaves Easter. At the end of the protease cascade, Easter cleaves Spätzle (Morisato and Anderson, 1994). This reaction apparently occurs shortly after fertilization and only on the ventral side of the embryo. Cleaved Spätzle is the ligand for Toll (Hashimoto *et al.*, 1988), a receptor on the egg membrane. The uniformly distributed Toll is thus only activated in the ventral side. Dorsal is held in the egg cytoplasm by Cactus. An entire pathway is then "designed" to separate Cactus from Dorsal in the ventral region. Toll signaling activates Pelle (protein kinase) and tube (function unknown yet). Pelle phosphorylates Cactus. Phosphorylated Cactus is degraded and Dorsal is free to enter the nucleus. This regulatory process leads to the high nucleus gradient in the ventral side and low gradient in the dorsal side. Fate of the cells is determined by the gradient of nucleus Dorsal protein. Genes that have a low affinity Dorsal binding region are

activated. Genes that have a high affinity Dorsal binding region are inhibited in conjunction with other enhancer regions. Target genes also influence one another. All of these finally lead to the formation of dorsal and ventral tissues.

The *windbeutel* transcript is present in ovaries, early embryos (0-4 hr) at the time when the dorsalventral pattern is established and in adult males but is almost undetectable in female carcasses. Careful examination of the wind expression pattern revealed that only the follicle cells that are located over the oocyte express the wind transcript, in contrast to the follicle cells covering the nurse cells (Konsolaki and Schupbach, 1998). In female windbeutel mutants, Pipe localization is altered from the Golgi to the ER. This was further confirmed by a heterologous expression system, using COS7 monkey cells (Sen *et al.*, 2000). When Pipe alone is expressed in COS7 cells, it clearly exhibits an ER distribution. However, when Wind is co-expressed, Pipe is translocated to the Golgi. These results clearly show that Wind is required for and enables the correct subcellular distribution of Pipe to facilitate its pattern-forming activity. The protein, Wind, like its mammalian homologue ERp28/29 is an ER luminal protein and belongs to the PDI-D subfamily. It contains 257 residues including a putative signal sequence at the N-terminus, and a KEEL ER-localization sequence at the C-terminus. The similarity between Wind and its mammalian homologue ERp28/29 is about 30%, however, both Wind and ERp28 show a similar domain organization, a N-terminal redox inactive thioredoxin domain and a C-terminal D-domain. Although Wind has a –CTGC- sequence at the very N-terminus, the position is obviously different from that of the –CXXC-motif in the redox active thioredoxin domain of other members of the PDI family. The requirement of Wind for the proper localization of Pipe suggests that Wind may function as a chaperone for the folding of Pipe in the ER. It could also be an escort protein to accompany Pipe from the ER to the Golgi like RAP (receptor associated protein) for the LDL receptor (Bu *et al.*, 1995; Bu and Schwartz, 1998; Bu, 2001).

3. Aims of the work

First, although some data support the idea that ERp28 may function as a putative chaperone, and possibly plays a role in the development process as its *Drosophila* homologue Wind does, we are still far from a clear understanding of the functions of ERp28, particularly *in vivo*. To gain further insight, ERp28 knock out mice would be very useful, especially as approaches using antisense technology (small interfering RNA) have so far proved unsuccessful (D.M.Ferrari, unpublished data).

The aim of the first part of my work is to generate ERp28 knock out mice. For this purpose, first, a mouse genomic DNA library will be screened with [³²P] labelled cDNA probe to obtain positive clones that contain the mouse ERp28 genomic locus. Then an ERp28 knock out vector will be made and transfected into ES cells. Positive homologous recombinant ES clones will be introduced into mouse blastocysts by microinjection to generate chimeric mice. Male chimeric mice are used to produce heterozygotes by mating with wild type female mice. These mice will be used for further breeding to generate -/- homozygotes. Phenotypic and functional analysis will be then carried out on these mice to clarify the function of ERp28.

Second, it is known that Wind is important for the dorsal-ventral patterning of the *Drosophila* embryo and it is clear that Wind is required for the proper localization of Pipe, another crucial protein for the dorsal-ventral patterning, to the Golgi. However, many questions still remain to be answered. For example, what are the concrete actions of Wind in the transport of Pipe? Is it a chaperone just for Pipe folding or an escort chaperone for the transport or both? How does Wind fulfill its functions to fold or transport Pipe? Is it necessary for other proteins to participate in the process of the folding or transporting of Pipe? A 3-dimensional structure would be very important not only for understanding the functions of Wind but also for other members of PDI-D family (for example, mammalian ERp28) and even for other PDI-related proteins. The aim of the second part of my work, therefore, is first to express and purify the His-tagged recombinant protein of Wind in *E.coli* and then crystallization of the purified protein will be attempted in order to solve the 3-dimensional structure of the protein. In addition, biochemical experiments and mutagenesis work will be performed to elucidate the function of Wind.

4. Materials

Chemicals

All chemicals were, if not mentioned elsewhere, analytical grade and were purchased from Boehringer, Merk, Sigma, Serva, Gibco BRL, New England Biolabs (NEB), etc; [α - 32 P]-dCTP was bought from Amersham Pharmacia, Hartmann;

Restriction Enzymes, T4 ligase, CIP and DNA polymerase

Restriction Enzymes:	NEB, MBI, and Roche
T4 DNA ligase:	MBI and NEB
CIP:	NEB
Taq DNA polymerase:	Roche
Pfu DNA polymerase:	Stratagen, Promega
TakRaLa-Taq:	TaKaRa

Bacterial strains

DH5 α :	Gibco-BRL
XL1 Blue:	Stratagene
BL21:	Stratagene

Plasmids

PQE30, PQE60:	Qiagen
PGEX-6p-2:	Pharmacia
pBluescript II:	Stratagene
TOPO TA cloning vectors:	Invitrogen
pGEM-T-easy vector:	Promega

Other plasmids used for making the ERp28 knock out vector were mentioned in the methods and/or results sections

Mouse strains

C57BL6N, NMRII and S129 were obtained from and raised in the Animal Facility of Max-Planck-Institute of Biophysical Chemistry.

Kits

Random primers DNA labelling system:	Invitrogen
Rediprime TM II random prime labelling system:	Amersham
Qiagen Spin Miniprep, Midiprep, Maximum kit	Qiagen
QIAquick Gel Extraction kit	Qiagen
PCR purification kit	Qiagen
Nucleotide Removal kit	Qiagen
The SuperSignal Substrate kit	Pierce

Membranes, Ni-NTA agarose, X-ray films

Nitrocellulose membrane	Sartorius
QIABRANE nylon membranes	Qiagen
Ni-NTA agarose	Qiagen
GST-Agarose beads	Sigma
X-ray film	Kodak

Primers

All of the primers were synthesized by MWG-Biotech AG

Frequently used medium, buffers and solutions

LB medium:	10g Bacotryptone, 5 g Yeast extract, 5 g NaCl, and 1 litre with distilled water
PBS (phosphate buffered saline):	137 mM NaCl 2.7 mM KCl 8.1 mM Na ₂ HPO ₄ 1.4 mM KH ₂ PO ₄ (pH 7.4)
TE (Tris/EDTA):	10 mM Tris-Cl, pH 7.5 1 mM EDTA
TBS (Tris-buffered saline)	0.9% NaCl (w/v) 50 mM Tris, pH 7.4

5. Methods

5.1 General methods:

5.1.1 PCR (polymerase chain reaction)

5.1.1.1 For DNA amplification from plasmid templates and short fragments from genomic DNA, the following reaction mixtures and conditions were used. The annealing temperature, extension time and the number of cycles were adjusted according to primers used and the size of the product expected.

The reaction mixture:

Template DNA	1 μ l (10-100 ng)	Typical reaction programme used:	
10 \times buffer	5 μ l	Temperature	Time
2 mM dNTP	5 μ l	94°C	5 min
forward primer	1 μ l (10 pmol/ μ l)	94°C	1 sec
reverse primer	1 μ l (10 pmol/ μ l)	55-60°C	1 min
DNA polymerase (Taq; Pfu)	1 μ l (2.5 u/ μ l)	72°C	2 sec
ddH ₂ O	36 μ l	repeat	30-35 cycles
		72°C	10 min
Total volume	50 μ l		

5.1.1.2 For long PCR amplification from genomic DNA, the reaction below was used:

Genomic DNA	2 μ l (200 ng-1 μ g)	Typical reaction program used:	
10 \times buffer	5 μ l	98°C	20 sec
2.5 mM dNTP	8 μ l	68°C	10-15 min
forward primer	1 μ l (10 pmol/ μ l)	repeat	35 cycles
reverse primer	1 μ l (10 pmol/ μ l)		
TaKa Ra LA Taq	1 μ l (2.5 u)		
ddH ₂ O	32 μ l		
Total volume	50 μ l		

5.1.2 Purification of PCR products

PCR products were run on an agarose gel, the desired bands were excised and purified by using a QIAGEN PCR purification kit according to the manufacturers instructions. PCR products could be purified directly if it was not necessary to separate the PCR products on the agarose gel.

5.1.3 DNA extraction from agarose gel

DNA was isolated using a QIAGEN gel extraction kit according to the manufacturers recommendations.

5.1.4 Estimation of DNA purity and quantitation

Quantification of DNA was performed on a Gene Quant II (Pharmacia) photometer. Both DNA and protein absorb light of wavelength 260 nm. However, proteins (aromatic side chains) absorb much stronger at 280 nm. Therefore, the ratio A₂₆₀/A₂₈₀ gives an estimation of the degree of contamination of a DNA sample with protein. For pure DNA, the A₂₆₀/A₂₈₀ is about 1.8. For DNA quantitation, the approximate relationship was used that an A₂₆₀ of 1.0 corresponds to ca. 50 µg/ml double-stranded DNA, 40 µg/ml single-stranded DNA and 31 µg/ml oligonucleotide DNA

5.1.5 Plasmid extraction

Small-, medium- and large-scale plasmid extractions were performed using plasmid mini/midi- and maxi-prep kits from QIAGEN according to the manufacturers instructions.

5.1.6 Restriction enzyme digestion, dephosphorylation, preparation of blunt-ended DNA and ligation

Restriction enzyme digestion was carried out according to a standard procedure. The restriction enzyme solution was kept below 10% total volume to prevent inactivation and star activity. Digested DNA was purified either by gel electrophoresis and extraction with a QIAGEN gel extraction kit or directly with a PCR purification kit.

To avoid undesired self ligation of plasmid DNA, the 5' phosphate groups of the vector DNA were dephosphorylated using calf intestinal phosphatase (CIP, NEB), employed at a concentration of 0.5u/µg DNA at 37°C for 60 min using the appropriate buffer delivered with the enzyme. Dephosphorylated DNA was purified by gel electrophoresis and gel extraction.

For ligations of incompatible ends of DNA, blunt-ended DNA was first prepared using the Klenow fragment of DNA polymerase according to standard procedure, followed by purification with the QIAGEN PCR purification kit.

T4 DNA ligase catalyses the ATP-dependent ligation of blunt or complementary sticky ends of DNA. The enzyme was used according to the supplier's instruction. Sticky-end ligations were carried out at room temperature for 2-4 hours or overnight, typically a 1:1 to 5:1 or higher insert:vector molar ratio and roughly 50-500 ng insert. Blunt-ended ligations were carried out at 14°C or room temperature overnight, with a 5:1 or even higher molar ratio of insert:vector.

Typical reaction mixture:

insert DNA	50-500 ng
Vector DNA	(about 100 ng)
1×ligase buffer	(including ATP)
T4 DNA ligase	1-2.5 U
ddH ₂ O	
total	20µl

5.1.7. Transformation of *E. coli*

5.1.7.1 Preparation of competent *E. coli* for heat shock transformation

A single colony of *E. coli* cells (XL1 Blue, DH5α) was inoculated into 50 ml LB medium and cultured overnight at 37°C with shaking (250 rpm). 4 ml of the overnight culture was transferred into 400 ml LB medium in a 2-liter flask, and grown at 37°C, with shaking (200 rpm), to an OD₆₀₀ of 0.375. The culture was aliquoted into eight 50-ml prechilled, sterile polypropylene tubes and the tubes were incubated on ice 5 to 10 min. The cells were centrifuged 7 min at 1600 × g, 4°C and each pellet was resuspended (gently) in 10 ml ice-cold CaCl₂ solution. The samples were spun down at 1100 × g for 5 min at 4°C and the pellet was again resuspended in 10 ml cold CaCl₂ solution. Then the cells were incubated on ice for 30 min and spun down at 1100 × g, 5 min, 4°C. Each pellet was resuspended completely in 2 ml of ice-cold CaCl₂ solution. The cells were aliquoted 250 µl each into prechilled, sterile polypropylene tubes. The cells were frozen immediately in liquid nitrogen and stored at -80°C.

CaCl₂ solution:	60 mM CaCl ₂
	15% (v/v) glycerol
	10 mM PIPES, pH 7.0

5.1.7.2 Preparation of electro-competent *E. coli*

A 10 ml overnight culture of *E. coli* (XL1 Blue, DH5 α) cells was used to inoculate in 1 l of fresh LB medium. The culture was grown at 37°C with agitation for about 2.5 h until OD₆₀₀ of 0.5-0.7 was achieved. The flask was cooled on ice for about 30 min, and the cells were collected by centrifugation at 4000 g, 4°C, for 20 min. The cells were resuspended in 1 l of ice-cold, sterile HEPES buffer, spun again as above, resuspended a second time in 500 ml HEPES buffer, spun, resuspended a third time in 20 ml HEPES buffer with 10% (v/v) glycerol, centrifuged again and finally resuspended in 2-3 ml sterile 10% (v/v) glycerol. The cells were then dispensed in 50-100 μ l aliquots and frozen on dry ice or liquid nitrogen. The frozen samples were stored at -80°C.

HEPES buffer: 1mM HEPES (pH 7.0) **HEPES/glycerol:** 1mM HEPES (pH7.0)

10% (v/v) glycerol

10% glycerol: 10% (v/v) glycerol

in distilled water

All solutions were sterilized by autoclaving before use.

5.1.7.3 Transformation by heat shock

50-100 μ l of competent *E. coli* were thawed on ice and added to 10-50 ng of plasmid mixture or to the ligation reaction. Cells were incubated for 30 min on ice, then heat shocked for 1.5 min (XL1 Blue) or for 45 sec (DH5 α) at 42°C in a water bath. The cells were then placed on ice for 1-2 min. Thereafter, 1 ml LB was added and the samples were incubated at 37°C for 45 min to allow initiation of the expression of selection marker proteins. 100-200 μ l was plated on agar plates containing selection antibiotics and incubated overnight at 37°C.

5.1.7.4 Transformation by electroporation

50 μ l of the electro-competent cells were thawed and 1-10 ng of salt free plasmid or 1 μ l of ligation reaction was added to the cells. The samples were mixed and transferred to a chilled 0.2 cm electroporation cuvette. Then the cuvette was transferred to a Gene Pulser electroporation chamber (Bio-Rad) and pulsed once with 25 μ F, 2500 V, 200 ohms. The time constant was usually about 4.5-5 ms. 1 ml LB medium was added immediately after the pulse and the sample transferred to a 1.5 ml Eppendorf tube and incubated with agitation at 37°C for 45min-1h. 10-100 μ l of the samples was then plated onto LB plates containing the selection antibiotic and incubated overnight at 37°C.

5.1.8 DNA gel electrophoresis

At neutral pH, the negatively charged DNA molecules migrate under the influence of an electrical field from the cathode to the anode. The distance migrated depends on fragment size and is fairly independent of base or sequence composition. The bands are made visible under ultraviolet light of 302 nm by staining the gel in an ethidium bromide solution. Alternatively, ethidium bromide can be included in the gel (0.5 µg / ml).

10×TBE	500 mM Tris	5×Blue marker:	20% (w/v) Ficoll
	10 mM EDTA		0.125% (w/v) Bromophenol blue
	500 mM H ₃ BO ₃		0.125% (w/v) Xylene-cyanol
	pH 8.0		

Ethidium bromide stock solution: 10 mg/ml

Agarose gel powder was added to an appropriate volume of 1 × TBE and boiled in a microwave oven until the agarose had dissolved. The liquid was then allowed to cool to ~50°C, whereupon the gel was cast. To increase the density of the applied sample and to allow monitoring of the progress of the electrophoresis, 0.2 volume blue marker was added to the samples. All gels were run in TBE at ~15 V/cm. Documentation was performed with the aid of a CCD camera, after staining for 10-15 min in ethidium bromide solution.

5.1.9 Protein gel electrophoresis

5.1.9.1 Polyacrylamide gel electrophoresis (PAGE)

The principle of polyacrylamide gel electrophoresis is separation of proteins of varying molecular weights under the influence of an electrical field in a continuous, cross-linked polymer matrix from the cathode to anode. Here, the polymer is polyacrylamide and the cross-linking agent bis-acrylamide. Cross-linking is effected through a radical-induced pathway by the addition of ammonium peroxide and TEMED (1,2-Bis (dimethylamino)-ethane).

Two gels are employed: a stacking gel with a low level of cross-linkage and low pH, allowing protein bands to enter the gel and collect without smearing, and a separating gel with a higher pH, in which the proteins are separated on the basis of size.

5.1.9.2 SDS-PAGE gel

The detergent sodium dodecyl sulphate is known to bind to proteins at an average of one SDS molecule per two amino acid residues, or roughly 1.4 g per g proteins, when present

at concentrations above 0.8 mM SDS is employed to effect denaturation of the proteins, to dissociate protein complexes and to impart upon the polypeptide chains net negative charge densities proportional to the length of the molecule. A reducing agent such as dithiothreitol (DTT) or β -mercaptoethanol (2-ME) is used to reduce any existing cystines (disulfide bonds). The electrophoretic mobility of such SDS-protein complexes is inversely proportional to the logarithm of the molecular mass of the protein. For an 8×6×0.1 cm gel, the following volumes were used:

	5% gels (Stacking)	10% gels (Resolving)	12.5% gels (Resolving)	15% gels (Resolving)
Solution A	-	1.5 ml	1.5 ml	1.5 ml
Solution B	625 μ l	-	-	-
Solution C	375 μ l	2.0 ml	2.5 ml	3.0 ml
H ₂ O	1.25 ml	2.5 ml	2.0 ml	1.5 ml
TEMED	3 μ l	3 μ l	3 μ l	3 μ l
APS (15%)	60 μ l	60 μ l	60 μ l	60 μ l

Solution A: 0.4% (w/v) SDS
1.5 M Tris-Cl pH 8.9

Solution C: 29.2% (w/v) acrylamide
0.8% (w/v) bis-acrylamide

Solution B: 0.4% (w/v) SDS
(stacking gel 0.5 M Tris-Cl pH 8.9
buffer)

Running buffer: 25 mM Tris-base
0.1% (w/v) SDS
192 mM glycine pH 8.75

APS: 15% (w/v)
Ammonium persulphate

Sample loading: 62.5mM Tris-Cl pH 6.8
buffer 3% (w/v) SDS
10% (v/v) glycerol
10% (v/v) 2-ME
0.0001% bromophenol blue

Solutions for separating gels were mixed according to the above table and poured immediately after addition of TEMED and APS, up to within 2 cm of the top of the gel cassette. A few drops of isopropanol were added to prevent dessication of the gel. After polymerisation was complete, the isopropanol was rinsed off and a freshly prepared stacking gel solution added. A constant 20 mA current was applied for electrophoresis of small

samples. Large sample volumes were run first at 15 mA until the bromophenol blue bands reached the separating gel, then at 20 mA until the bands reached the bottom of the gel.

5.1.9.3 Native gel (non-denaturing gel)

Recombinant His- wind and Wind-His protein were run on 7.5% native gels, using samples treated with 100 mM DTT was added to the samples.

The basis of separation of native gel or non-denaturing gel is as described in the section 5.1.9.2, with the exception that proteins now separate according to an apparent molecular weight based on the overall size and shape of the molecule. SDS is not added to such gels, and they are run at lower voltages and temperatures to reduce the risk of heat-denaturation of the proteins. In this work, the gel solutions used were identical to those used for SDS-PAGE, with the exception that SDS was omitted from all the solutions. 2-ME or DTT was still present to preserve the reduced state of the proteins.

Sample loading buffer: 62.5 mM Tris/HCl
(non-denaturing gel) 10% (v/v) glycerol
0.001% (w/v) bromophenol blue

7.5% native gel

Solution A: 1.5 M Tris-Cl (pH 8.9) 1.5ml
Solution C: 29.2% acrylamide
0.8% bis-acrylamide
water: 3 ml
TEMED: 3 µl
40% APS: 22.5 µl

Loading buffer

62.5 mM Tris-Cl, pH 6.8
10% glycerol
10% 2-ME
0.001% bromophenol blue

5.1.10 Coomassie brilliant blue staining

Gels were stained in staining solution for 0.5-1 h at room temperature followed by reduction of the background using destaining solution for several hours at room temperature.

Staining solution: 0.5% (w/v) Coomassie brilliant blue R 250
45% (v/v) ethanol
10% (v/v) acetic acid

Destaining solution: 45% ethanol
10% acetic acid

5.1.11. Western blotting

5.1.11.1 Transferring of proteins onto nitrocellulose membranes

Blot buffer: 25 mM Tris-Cl; pH 8.3
 192 mM glycine
 20% (v/v) methanol

A semi-dry set-up (Millipore Milli-Blot SDE-system) was employed. Upon completion of electrophoresis, the polyacrylamide gel was removed and placed on top of two sheets of Whatman filter paper and a nitrocellulose membrane (0.2 μ m pore diameter), cut to size and pre-equilibrated in Blot buffer. After carefully removing any air bubbles present, a further two sheets of Whatman filter paper, similarly pre-equilibrated, were applied. The transfer sandwich, which was set up on the cathodal plate of the chamber, was covered with a lid comprising the anode. For an 8×6×0.1 cm gel, transfer was carried out at 100 mA constant current for 1 hour.

5.1.11.2 Antibody staining of Western blots

Ponceau S	0.1 % (v/v) Ponceau S	antibody incubation:	0.5% (w/v) low fat
solution:	5% acetic acid	buffer	milk powder in PBS

Blocking solution: 5% (w/v) low fat milk powder in PBS

PBS (phosphate buffered saline): 137 mM NaCl
 2.7 mM KCl
 8.1 mM Na₂HPO₄
 1.4 mM KH₂PO₄, pH: 7.4

Briefly adding Ponceau S solution to the membrane allowed visualisation of the transferred bands. The blotted membrane was washed with PBS, treated with blocking solution for at least 2 h at RT or with antibody incubation buffer overnight at 4°C, then with blocking buffer 1-2 h at RT to saturate protein binding sites on the membrane and then incubated with primary antibody (generally at 1:1000 dilution) in 10 ml antibody incubation buffer for 1 h at RT. After three washes in PBS or antibody incubation buffer, the membrane was similarly incubated with the second, horseradish peroxidase-coupled antibody (diluted 1:2000-1:40,000 in incubation buffer).

5.1.11.3 Detection with peroxidase/diamino benzidine

Staining solution: 0.1% (w/v) CoCl_2
0.05% (w/v) diamino benzidine/HCl
0.2% (v/v) H_2O_2

Horseradish peroxidase converts 3,3'-diaminobenzidine into a brown precipitate in the presence of its substrate peroxide. Adding nickel or cobalt ions enhances the sensitivity of the reaction. Here, the secondary antibody was used at 1:2000 dilution.

After three or four washes in PBS or PBST (PBS with 0.05% Tween-20) for 10 min. the blocked nitrocellulose membranes were stained for up to 1 min using 10 ml freshly prepared staining solution. The staining solution was removed by washing in PBS, and the blots were stored air-dried.

5.1.11.4 Chemiluminescence

Detection by chemiluminescence was performed using the SuperSignal Substrate kit (Pierce) as recommended by the manufacturer. The signal was recorded on Cronex 4 X-ray film (DuPont) or on a Fujifilm LAS-1000 cooled CCD instrument using Image Reader and Raytest Aida image analysis software (Fuji). Preparation of the blot was essentially as in section 11.3, apart from that HRP-coupled secondary antibody was used at a concentration of 1:10 000-1:40 000 and that Tween-20 was added to the TBS or PBS wash buffers to a final concentration of 0.05%. Because of the high sensitivity of the method, at least four wash steps of 15 min each were performed before application of the detection solution.

5.1.12 Protein quantitation

Protein concentration were estimated according to the method of Bradford (Bradford, 1976). Briefly, the absorption of a protein-specific dye, Coomassie Brilliant Blue, is measured near 595 nm. This absorption, which is due to interactions with basic or aromatic residues, yields a relatively good linear concentration dependence for most soluble proteins. Protein solutions of unknown concentration are measured and the absorbance plotted against a reference curve obtained with a protein of known concentration. From this an approximate concentration is determined.

Dye solution: 100 mg	Coomassie Brilliant Blue G250
50 ml	100% ethanol
100 ml	85% (v/v) phosphoric acid
Ad 1000 ml	with double distilled water

The dye was first dissolved in ethanol before addition of the other components. The solution was filtered before use. A few μ l of the protein sample were added to 1 ml dye solution. The mixture was incubated for 5 min at room temperature and then measured spectrophotometrically at 595 nm or 578 nm. As reference, an identical amount of protein solvent solution was treated similarly. The absorbance value was compared to that of a BSA standard curve to obtain protein concentrations.

5.2 Generation of knock out mice

5.2.1 A brief introduction of the knock out strategy and the knock out vector

Gene knock out in mice is a powerful tool for studying the functions of a gene *in vivo*. The aim of this approach is to delete a gene of interest in mice, thus causing a loss of the functions of the gene. By analyzing the phenotype of the knock out mice, the functions of the gene of interest can be clarified or deduced. Knock out is based on and is the most popular usage of gene targeting. Gene targeting, defined as the introduction of site-specific modifications into the mouse genome by homologous recombination, is generally used for the production of mutant animals to study gene function *in vivo*. Since homologous recombination of foreign DNA with endogenous genomic sequences is a relatively infrequent event in mammalian cells, compared to random integration, the only efficient gene targeting method presently established utilizes pluripotent murine embryonic stem cell lines. Using these cells, the selection of rare, homologous recombinant ES cell clones *in vitro* can be accomplished. When such genetically modified ES cells are introduced into a preimplantation embryo they can contribute, even after extensive *in vitro* manipulation, to all cell lineages (including germ cells) of the resulting chimeric animal. The breeding of germline chimeras, which transmits an ES cell derived mutant chromosome(s) to their progeny, allows the establishment of an animal heterozygous for the derived genetic alteration and, more importantly, by further breeding a homozygous mutant mouse strain. The necessity of chimera production may be bypassed in the future when spermatogonial stem cell lines become available.

As a substrate for homologous recombination, vectors of the replacement type (here referred as knock out vector) are most frequently used for gene targeting in ES cells and are intended to simply inactivate gene function. A typical replacement vector consists of two regions of DNA (4-10 kb in total) homologous to the genomic target locus which are interrupted by a positive selection marker such as the bacterial aminoglycoside phosphotransferase (neo) gene which is selected for with the potent ribosome binding drug,

G418. The positive marker may serve two functions. Its primary purpose is as a selection marker to isolate the rare transfected cells that have integrated DNA (which occur at a frequency of about one in 10^4). Secondly, it can serve as a mutagen, for instance if it is cloned into a coding exon of the gene or replaces coding exons. If recombinant ES clones will be identified by PCR, one of the vector arms must be kept relatively short (1-2 kb) to ensure efficient amplification. A thymidine kinase (tk) gene is often included at the end of the long homology arm of the vector to serve as an additional negative selection marker (using the drug ganciclovir) against ES clones which have randomly integrated the targeting vector. Thus, homologous recombinants can be enriched by both positive and negative selection.

In summary, the whole knock out procedure can be divided into the following steps: preparing the knock out vector; transfecting the ES cells, and screening the ES cells to get the positive recombinant ES cells; microinjecting the positive ES cells into the mouse blastocyst to generate chimeric mice; crossing the chimeric mice with wild type mice to get heterozygotic mice; recrossing the heterozygotic mice to produce homozygotic mice; finally, analysing the phenotype of knock out mice to work out the function of the knocked out gene.

5.2.2 Searching and alignment of the mouse ERp28 ESTs from the database

Using the cDNA sequence of human ERp28 to search the mouse EST database in Genbank, several mouse ESTs were found. By alignment and comparison of these ESTs using the programme DNASTAR, nearly full length mouse ERp28 cDNA including all of the coding region was obtained.

5.2.3 The exon distribution of mouse ERp28

The cDNA sequence of ERp28 obtained above was used to search the database and look for fragments of known mouse genomic DNA that contained pieces of mouse ERp28 cDNA. Comparing the cDNA with these genomic DNA fragments, the exon distribution of mouse ERp28 could be deduced. To find out the distance between the exons (intron regions), PCR was performed using primers (P278/P293, P294/P279) designed from different exons and using mouse genomic DNA as template.

P293 5' CTGCCAGACGCTTGAACCTCATCTTGCTTCTCTCCATAG 3'

P294 5' CTATGGAGAGAAGCAAGATGAGTTCAAGCGTCTGGCAG 3'

The sequences of P278 and P279 are shown in section 5.2.5.

5.2.4 RNA extraction

Total RNA was extracted from the brain of an E12.5 mouse embryo and cultured F9 cells using the RNeasy Mini Kit from Qiagen according to the protocol provided by the manufacturer.

5.2.5 RT-PCR to obtain a cDNA probe for mouse Genomic DNA library screening

Two primers for RT-PCR were designed from the sequences of the first and second exons, respectively (as shown in fig 6.1.2)

Upstream primer (P278) 5' TGCTCCTGGGCCTGCTGCTCCTCTTTGCTCCGCAC3';

Downstream primer (p279) 5' GCATCTCACCAGCCATCTTGACTGAGCACACAGGC 3'.

The template RNA used is total RNA extracted either from E12.5 mouse brain, or cultured teratocarcinoma F9 cells.

Reaction mixture:

	Volume	Final concentration
RNase-free water	17.0 µl	—
5x QIAGEN OneStep RT-PCR Buffer	10.0 µl	1x
dNTP Mix (containing 10 mM of each dNTP)	2.0 µl	400 µM of each dNTP
5x Q-Solution	10.0 µl	1x
P278 (upstream primer)	3 µl	0.6 µM
P279 (downstream primer)	3 µl	0.6 µM
QIAGEN OneStep RT-PCR Enzyme Mix	2.0 µl	—
RNase inhibitor (optional)	1.0 µl	5–10 units/reaction
Template RNA	2.0 µl	1– 2 µg/reaction
<hr/>		
Total volume	50.0 µl	—

Program used: 50°C, 30 min for reverse transcription; 95°C, 15 min to activate the Enzyme Mix and inactivate the reverse transcriptases; followed by PCR: 94°C, 30 sec; 60°C, 30 sec; 72°C, 2 min; 35 cycles in total and then 72°C, 10 min.

Products from the RT-PCR reaction were purified and cloned into TOPOII TA cloning vectors (Invitrogen).

5.2.6 TA cloning

The PCR products amplified by Taq DNA polymerase can be cloned into the TA cloning vectors due to an extra A at the 3' terminus of the PCR products. For high fidelity DNA polymerase like Pfu, the extra A can be added by incubating the products with Taq polymerase for 10-20 minutes at 72°C after the PCR reaction. The TA cloning vectors were used according to the instructions of the manufacturers (TOPO TA cloning vector from Invitrogen; pGEM-T-Easy TA cloning vector from Promega).

5.2.7 Genomic DNA library screening

5.2.7.1 Library titre determination (mouse S129 genomic DNA library in Lambda Fix II, Stratagene, kindly provided by Professor Nils Brose, Max-Planck-Institute of Experimental Medicine, Göttingen).

1. Preparation of bacterial culture for infection (XL1-MRA)

A single colony of XL1-MRA was picked and put into 3 ml LB with 0.2% maltose and 10 mM MgSO₄, cultured at 37°C, 180 rpm, until an OD₆₀₀ of 0.6-0.8 (about 5 hours).

2. Titration of packaged phages on LB plates

Phage buffer: 20 mM Tris-Cl, pH 7.4
100 mM NaCl
10 mM MgSO₄

A series of dilutions were made in phage buffer, 1:10³, 1:10⁴, 1:10⁵, 1:10⁶. 100 µl of the diluted phages was added to 100 µl of prepared bacterial cells in a 1.5 ml tube, mixed gently, incubated, 37°C, 30 min. The mixture was transferred to a 15 ml tube containing 6.5 ml prewarmed top agarose (45°C), mixed gently by inverting, and immediately poured on top of prewarmed LB plates (37°C). The plates were incubated at 37°C, overnight. The number of plaques on the plates was counted and the titre of the library was calculated.

5.2.7.2 The first round of screening

1. Preparation of the LB plates for screening

First, four big plates were prepared with 300-400 ml of LB agar, hardened and dried at 37°C overnight or room temperature for 2 days before using.

XL1-MRA host bacteria was prepared as before. 1 µl of the library was diluted to 1:10 and 1 µl hereof was used for each plate (corresponding to 2×10⁵ plaques per big plate, this contains almost a whole mouse genome according to the size). 1 µl of the diluted phages was added to 100 µl of prepared bacterial cells in a 1.5 ml tube, mixed gently and incubated,

at 37°C, for 30 min. The mixture was transferred a 50 ml tube containing 40 ml prewarmed top agarose (45°C), mixed gently by inverting, immediately poured on top of prewarmed LB plates (37°C). The plates were incubated at 37°C, overnight or until the plaques reached a suitable size without merging.

2. Transfer of phages to HybondTM-N membranes.

HybondTM-N membranes (two membranes for each plate) were placed onto the surface of plates with suitably sized plaques for 1 min. The membranes and plates were marked by puncturing with a needle to ensure correct orientation of colonies/plaques. After 1 min, the membranes were placed, colony side up, on a pad of absorbent filter paper soaked in denaturing solution, for 7 minutes. Then the membrane was placed, colony side up, on a pad of absorbent filter paper soaked in neutralizing solution, for 3 minutes, and then repeated with a fresh pad soaked in the same solution. The membranes were washed in 2×SSC, then air dried on the dry filter paper, colony side up. Finally the membranes were baked at 80°C for 2 hours to fix the DNA on the membranes.

Denaturing solution: 1.5 M NaCl
0.5 M NaOH

Neutralizing solution: 1.5 M NaCl
0.5 M Tris-Cl pH 7.2
0.001 M EDTA

20×SSC stock solution: 3 M NaCl
0.3 M sodium citrate

3. Prehybridization

Solutions:

20×SSPE: 3.6 M NaCl	5% SDS
0.2 M sodium phosphate	0.2 M EDTA (pH 8.0)
0.02 M EDTA pH 7.7	
100× Denhardt's Solution: 2% BSA (bovine serum albumin)	
2% Filcolin TM	
2% pvp (polyvinylpyrrolidone)	

Sonicated non-homologous salmon sperm DNA (10 mg/ml) was first denatured by heating at 95-100°C for 5 min, then chilled on ice and added to the solution prepared as below to a final concentration 0.02 ng/ml. 25 ml of prehybridization solution was added to each of

four hybridization bottles, to which the membranes were then added. The membranes were then prehybridized in a hybridization oven at 65°C, overnight.

Pre-hybridization solution was prepared as below

Solution	Volume	Final Concentration
20×SSPE	6.25 ml	5×
100×Denhardt's solution	1.25 ml	5×Denhardt's
5% SDS	2.5 ml	0.5% SDS
total	25 ml (for one membrane in one hybridization bottle)	

4. Hybridization

DNA Probe was labeled using the RedprimeTM_{II} Kit (random primer labeling system) (Amersham Pharmacia). The cDNA insert from the TOPOII vector was digested by EcoRI and purified by gel electrophoresis and gel extraction, 25 ng of insert DNA in 45 µl of TE buffer (10mM Tris-Cl, pH 8.0, 1 mM EDTA) was denatured by heating at 95-100°C for 5 min, cooled rapidly on ice for 5 min, then briefly centrifuged. The denatured DNA was transferred to the reaction tube, 5 µl (10 µCi/µl) of Redivue^[32P]dCTP (Amersham Pharmacia) was added and the solution was mixed by pipetting 12 times. After a 10 min incubation at 37°C, 5 µl of 0.2 M EDTA (pH 7.5-8.0) was added to stop the reaction. The labeled probe was denatured at 95-100°C for 5 min, then cooled on ice for 5 min and spun down briefly. One reaction mix was added to each of the hybridization bottles and hybridization was carried out overnight at 65 °C.

5. Wash

After hybridization, the membranes were washed with washing buffer (0.1×SSPE, 0.1% SDS) for 4 hours, changing the washing buffer every hour. The membranes were then wrapped in Saran wrap, and detected by autoradiography (exposed for 3.5 days, -80°C).

5.2.7.3 The second round of screening

The positive plaques were picked from the first round of screening using Pasteur pipettes, transferred to a 1.5 ml Eppendorf tube containing 500 µl of SM buffer, and incubated at 4°C overnight to release the phages from agarose.

SM buffer: 0.1% gelatin
50 mM Tris-Cl (pH 7.5)
100 mM MgSO₄

1 µl was used as template and PCR was performed to confirm the positive clones again and to check whether they included the full length of the ERp28 locus. Primer pairs used for PCR were P299/P279, P301/P293, P273/P323 (fig 6.1.7) designed from the known sequences of genomic DNA and the mouse ERp28 cDNA.

P299: 5' GGAGCTGAGTGAGAAGTACAAGCTGGACAAGG AGAGCTAC 3';

P279: 5'GCATCTCACCAGCCATCTTGACTGAGCACACAGGC 3';

P301: 5' GTGTTTACCGAGAAGGACCTCCTGAGATGGCAGACCTGTCTG 3';

P293: 5' CTGCCAGACGCTTGAAGTCACTTGTCTTCTCTCCATAG 3';

P273: 5'GCTGCGCC ATGGGGAGCAAAGAGGAGCAGCAGGCCC 3';

P323: 5'GGACGGGACTTAGATGTG GAAGGTACTAGTTCCAG 3'

The above solution was diluted to 1:10⁵ in phage buffer and 100 µl was transferred to 100 µl of XL1-MRA host bacteria, and incubated at 37 °C for 30 min. The mixture was added to 7 ml prewarmed top agarose, poured on top of the LB plates (15 cm), and incubated at 37°C, overnight.

Plates with 500-1000 plaques (plaques were clearly separated from each other) were chosen. Phages were transferred to membranes and the orientation of the membranes was marked as in the first round of screening.

The membranes were baked at 80°C, for 2 hours to fix the DNA on the membranes. Prehybridization, hybridization, wash, and autoradiography were carried out as previously described.

5.2.7.4 The third round of screening by PCR

Several single positive plaques from each of the plates from the second round of screening were picked and the phages were released from agarose as before. The phages were diluted to 1:10⁵ and fresh plaques were prepared (separated from each other) as before. Single

plaques were picked and the phages were released from of the agarose by soaking the plaques in 500µl of SM buffer, at 4°C, overnight. PCR was performed as in the second round of screening.

5.2.8 Isolation of DNA from the positive plaques

The Wizard[®] Lambda Prep DNA purification System (Promega) was used to purify Lambda phage DNA by the liquid culture method according to the instructions of the manufacturer.

5.2.9 Restriction mapping

The phage DNA was either digested by NotI alone or double-digested by NotI and EcoRI to isolate the ERp28 locus. The digested fragments that contained the fragments of the gene locus of mouse ERp28 were then purified and cloned into the vector pBluescript II (Stratagene). The vectors were further digested by different combinations of restriction enzymes. A restriction map of the ERp28 locus could be determined from the digestion results.

5.2.10 Steps for building the knock-out vector

This will be described in detail in the results section (section 6.1.7).

5.2.11 Linearisation of the knock out vector

The knock out vector was digested by NotI to linearise the vector completely (linearised vectors increase the efficiency of recombination). Then the linearised vector was precipitated with 0.3 M NaAc (pH 5.2) and 3 volumes of 100% ethanol. Samples were centrifuged at maximum speed for 30 min, and the supernatant was discarded carefully. The pellet was washed with 500 µl of 70% ethanol and centrifuged again for 5 min. The supernatant was decanted carefully and the DNA was air dried. The linearised vector was dissolved in a suitable amount of autoclaved ddH₂O and the concentration (usually about 1 µg/µl) was determined.

5.2.12 ES cell transfection and selection

ES cells (MPI-II, passage 11) and anti-G418 fibroblast cells (feeder cells) were obtained from the Department of Molecular Cell Biology, Max-Planck-Institute of Biophysical Chemistry, Göttingen.

Schedule for the ES-cell experiment**Day 1****Initiation of fibroblast feeder cell culture**

Fibroblast medium (FB) DMEM with 2 mM glutamine,
or Emfis medium: 10% serum and 0.9% (w/v) glucose.

9 ml of medium was added to a 50 ml Falcon tube. One vial of fibroblast cells was thawed, transferred to the tube, and then spun down at 1000 rpm for 10 min. The medium was aspirated and the pellet was resuspended in 10 ml fresh medium and pipetted up and down several times. The cells were transferred to a 15 cm dish which contained 14 ml fresh medium. The cells were mixed with medium and incubated at 37°C, 5% CO₂, overnight. Meanwhile, 3ml medium was added to a 3.5 cm dish as a control to monitor contamination.

Day 2-3

The medium was changed every day and the cells were allowed to grow until confluency.

Day 4

Subculture of feeder cells. The confluent feeder cells were split into 3 dishes. The cells were washed once with PBS, 5 ml trypsin was added and incubated at 37°C for 5 min. 10 ml of fresh medium was then added, and the cells were pipetted up and down to remove clumps. The cells were transferred to a sterile 50 ml Falcon tube, and spun down at 1000 rpm, for 5-10 min. The cells were resuspended in 5 ml of fresh medium and 5ml of the cells was transferred to each of three plates containing 15-20 ml of fresh medium (two of them would be inactivated and frozen for culturing ES cells, the third one for subculture on 6 new plates).

Trypsin solution: 8.0 g of NaCl, 0.4 g of KCl, 0.1 g of Na₂HPO₄·2H₂O, 1.0 g of glucose, 0.3 g of trizma base, 0.01 g of phenol red and 2.5g of trypsin were dissolved in H₂O, the pH was adjusted to 7.6 and H₂O was added to 1000 ml. The solution was filter-sterilized and kept in 10-ml aliquots at -20 °C. For use, the solution was diluted (1:4) with saline/EDTA and stored at -20 °C.

Saline/EDTA: 8.0 g of NaCl, 0.2 g of KCl, 1.15 g of Na₂HPO₄·2H₂O, 0.2 g of KH₂PO₄ and 0.2 g of EDTA were dissolved in distilled water, the pH was adjusted to 7.2 and H₂O was added to a total of volume of 1000 ml. The solution was autoclaved and stored at RT.

Day 5-8

Inactivation and freezing of fibroblast feeder cells. For 15 cm dishes, the old medium was aspirated, and 10 ml fresh medium containing 100 µg/ml of mitomycin C was added. The cells were incubated at 37°C for 2.5 hours to inactivate the feeder cells. Then the inactivation medium was aspirated and the cells were washed once with PBS. 5 ml trypsin was added and the cells were incubated at 37°C for 5 min. 10 ml of fresh medium was added and the cells were pipetted up and down to remove cell clumps. The cells were collected from the two dishes into a sterile 50 ml Falcon tube, and spun down at 1000 rpm for 5-10 min. The pellet was resuspended in 1.5 ml of fresh medium and 1.5 ml of freezing medium was added. The sample was mixed and the cells were aliquoted to 3 Eppendorf tubes (2 ml), 1 ml in each. The cells were first frozen at -20°C for 1h; then at -80°C, overnight; for long term storage, the cells were further frozen in liquid nitrogen and stored at -180°C to -200°C. The cells in the third dishes were split into 6 new big dishes as before and allowed to grow for two days until confluency, and inactivated as above.

Day 9

Preparation of gelatinized dishes. 3.5 cm dishes were gelatinized with 2 ml of 0.1% gelatin for at least 15-20 min at RT). The cells in big dishes were washed with PBS, then incubated with 5 ml trypsin at 37°C for 5 min. 10 ml of the fresh medium was added to a big dish of cells and pipetted up and down 10-20 times. The cells were transferred to a 50 ml Falcon tube, medium was added to 40 ml, and pipetted up and down. The gelatine was aspirated from the small dishes and 2 ml of the cells was added to each dish and incubated at 37°C overnight. Meanwhile, the cells of 5 other dishes were frozen for later usage as before.

Gelatin solution: 1 g of gelatine was dissolved into 1000 ml of H₂O, autoclaved and stored at 4°C.

Day 10-12

Start of ES cell culture. One vial of ES cells was thawed quickly in a 37°C water bath. The thawed ES cells were transferred to a 15 ml Falcon tube containing 5 ml fresh F.B medium, and spun down at 1000 rpm for 5-10 min. The cell pellet was resuspended in 2 ml ES medium. Meanwhile, the medium in one 3.5 cm dish of inactivated feeder cells was removed. The ES cells were transferred to the dish and incubated at 37°C for two days without changing the medium.

ES medium: DMEM containing 4.5 g/L glucose, 2 mM β -mercaptoethanol, 2 mM glutamine, 1% (v/v) stock solution of non essential amino acids, 1 mM sodium pyruvate, 15% (v/v) FCS, and 500 U/ml of LIF (leukemia inhibitor factor).

Day 13

Subculture of ES cells. ES cells were subcultured on 5 new 3.5 cm dishes containing the inactivated feeder cells. The ES cells were first washed with 2 ml PBS, 0.5 ml of trypsin was added and incubated at 37°C for 5 min. 1.5 ml of fresh ES medium was added to inhibit the activity of trypsin and the cells were pipetted up and down carefully to obtain single ES cells (important for preventing the differentiation of ES cells). The cells were transferred to a 50 ml Falcon tube and spun down. The cell pellet was resuspended in 10 ml of fresh ES medium and 2 ml was transferred to each of five 3.5 cm dishes. The ES cells were incubated at 37°C.

Day 14

Change of the ES cell medium and preparation of two 10 cm dishes of feeder cells for transferring ES cells for day 15. 5 ml gelatin was added to each of two 10 cm dishes and incubated at least 20 min at RT. One tube of inactivated fibroblast feeder cells prepared before was thawed, and the cells were transferred to a 50 ml Falcon tube containing 9 ml of fresh feeder medium, and spun down at 1000 rpm for 5-10 min. The cell pellet was resuspended in 20 ml fresh feeder medium and 10 ml was transferred to each of the two gelatinized dishes.

Day 15

Transfer of the ES cells in small dishes to the two 10 cm dishes containing feeder cells. The medium was aspirated from the small dishes and the cells were washed with 2 ml PBS. 0.5 ml trypsin was added to each of the 3.5 cm dishes, incubated at 37°C for 5 min. Then 1.5 ml fresh ES medium was added and the cells were pipetted up and down carefully to obtain single ES cells. The cells were transferred to a Falcon tube, spun, and the pellet was resuspended in 10 ml of fresh ES medium. 5 ml was transferred to each of 10 cm dishes containing 5 ml of fresh medium. The cells were incubated at 37°C.

Day 16

Change of the medium of ES cells and preparation of 2 × 15 cm dishes of feeder cells for electroporation. 2 × 5 cm dishes were gelatinized. Two tubes of fibroblast cells were thawed in a 37°C water bath, transferred to a Falcon tube containing 8 ml of feeder medium, and spun

down. The cell pellet was resuspended in 10 ml of fresh medium. 5 ml was transferred to each of the two 15 cm dishes containing 15 ml medium. The dishes were incubated at 37°C.

Day 17

Transfection of ES cells by electroporation. The feeder medium was aspirated from the two 15 cm feeder cell dishes and exchanged for 25 ml ES medium. The ES cells in the two 10 cm dishes were trypsinized by adding 2 ml trypsin to each dish, incubated at 37°C for 5 min, and then 6 ml of PBS (note: not medium) was added. The cells were pipetted up and down to obtain single cells and then transferred into a Falcon tube, spun down and the cell pellet was resuspended in 0.8 ml of PBS. 30 µg of linearized knock-out vector was added and mixed with approximately 1.5×10^7 ES cells. The mixture was transferred to an electroporation cuvette (0.4 cm). Electroporation was performed under the following conditions: 250 V, 500 µF, time constant about 7 ms at room temperature. The electroporated cells were left at RT for 5-10 min, then transferred to 2×15 cm dishes containing feeder cells, and incubated at 37°C.

Day 18-25

Selection by Gancyclovir and G418

24 hours after electroporation, selection was started by changing the ES medium in the dishes with ES selection medium (25 ml). The ES cells were selected for 7-8 days (the last two days only selected by G418), changing the selection medium every day.

Selection medium: 200 µl of 2 mM Gancyclovir and 200 µl of 250 mg/ml G418 were added to 200 ml ES medium.

Day 24

Change of the selection medium (ES medium only with G418) on the sixth day of selection.

Preparation of some gelatinized 24-well plates (100-200 µl of 0.1% gelatin was added to each well, and the plates were incubated at 37°C overnight).

Day 25

Change of the selection medium (ES medium only with G418) on the seventh day of selection.

Preparation of 5× 24-well plates of feeder cells: 2 tubes of the frozen inactivated feeder cells were thawed at 37°C, transferred to a Falcon tube containing 8 ml of fibroblast cell medium (FB medium), and spun down at 1000 rpm for 5-10 min. The cells were resuspended in 15 ml of feeder cell medium, and 5 ml of the cells was transferred to each of three 50 ml Falcon tubes. More medium was added to 40.5 ml. Gelatin was aspirated from the 24-well plates, and 1ml of the cells was added to each well. Then the cells were incubated at 37°C.

Day 26-28

Picking of surviving clones. First, the clones that were big enough were marked, then the FB medium in the 24-well plates was replaced by 1 ml of ES medium. 30 µl of trypsin was added to each well of 96-well plates, the marked colonies were picked under a microscope with 10-µl micropipette tips, and transferred to 96-well plates. After 5-10 min of incubation at RT, 50 µl of ES medium was added to inhibit the activity of trypsin using a 12-channel micropipette. The cells were disaggregated by pipetting up and down and transferred to 24-well plates containing the inactivated fibroblast feeder cells in order.

Day 28

Splitting of some clones in the 24-well plates picked on the first day. The cells in the wells were split if they contained only few clones which were very big and compact, because they were derived from clumps of ES cells and not from single cells. These ES cells must be split, otherwise, they will tend to differentiate. If there were many ES cells in the well and the ES cells were flat and similar in size, they were allowed to grow until confluency.

Day 29-42

Freezing of the surviving ES cells and subculture for genomic DNA extraction. The cells were washed once with PBS, then trypsinized for 5 min with 100 µl of trypsin. 500 µl of fresh medium with G418 was added, the cells were pipetted up and down and 490 µl was transferred to freezing tubes marked according to the number of the wells, then 500 µl of 2× freezing medium was added to the freezing tubes. The rest of the ES cells was transferred to gelatinized 24-well plates (each well containing 1 ml ES medium with G418), and cultured at 37°C until confluency for genomic DNA extraction. The freezing tubes were first placed at -20°C overnight, then transferred to -80°C. For long term storage the cells were kept at -180°C.

2×Freezing medium (10 ml):	1 ml FCS
	1 ml DMSO (dimethylsulfoxide)
	8 ml ES medium

When the cells in 24 wells were confluent (usually after 2-3days), the cells were washed once with PBS, then 500 µl of lysis buffer was added, incubated at 37°C, overnight (at least 6 hours). The next day, the solution was transferred from each well to a 1.5 ml Eppendorf tube for DNA extraction.

Lysis buffer: 100 mM Tris-Cl, pH 8.5
5 mM EDTA
0.2% (W/V) SDS
200 mM NaCl
100 µg/ml Proteinase K

5.2.13 Extraction of genomic DNA from the surviving ES cells

500 µl of isopropanol was added to the Eppendorf tubes from step 5.2.12 (day 29-42) to precipitate the genomic DNA. The solution was well mixed and allowed to stand for a few minutes at RT, centrifuged at 13,000 g for 10 min at RT, and the supernatants were carefully discarded. The DNA pellets were washed with 500 µl of 70% ethanol (v/v), and centrifuged at 13,000 g for 5 min at RT. The supernatants were discarded and the pellets air dried. Each pellet was dissolved in 100 µl of ddH₂O (or 1 × TE, pH 8.0) and incubated at 37°C overnight.

5.2.14 Screening of the recombinant ES cells by PCR

Primers used are shown in fig 6.1.12 of the results Section. P366/LacZ-3 are external primers designed from the region outside 5' end of the left arm and the 5' end of the sequence of LacZ, respectively; P373/P374 are external primers designed according to the 3' end of the gene for neomycin resistance and the region outside 3' of the right arm, respectively. Since both pairs of primers are not included in the arms, they can be used to distinguish homologous recombination from random recombination. These primers were also used for genotyping of chimeric, heterozygotic, and homozygotic mice.

P366: 5' CTGGCAACACTCCAATGGAGGCCTCTGGTG 3'

LacZ-3: 5' AATGCCTCCCAGACCGGCAACGAAAATCACG 3'

P373: 5' CATGGCGATGCCTGCTTGCCGAATATCATGG 3'

P374: 5' CTGTCAGCTGACCATGGTCCGTATGGACTGC 3'

The reaction mixture:

Left fragment		Right fragment	
DNA	2µl (200ng)	DNA	2µl (200ng)
10×Buffer	5µl	10×Buffer	5µl
2.5mM dNTP	8µl	2.5mM dNTP	8µl
P366	1µl (10pmol/µl)	P373	1µl (10pmol/µl)
LacZ-3	1µl (10pmol/µl)	P374	1µl (10pmol/µl)
TaKa Ra LA Taq	0.5µl (5u/µl)	TaKa Ra LA Taq	0.5µl (5u/µl)
ddH ₂ O	32.5µl	ddH ₂ O	32.5µl
Total	50µl	Total	50µl

PCR programme used: 98°C 20 sec
68°C 15 min
repeat 35 cycles
72°C 10 min

5.2.15 Preparation of positive ES cells for blastocyst injection**Several days (3-4 days) before blastocyst injection**

The positive recombinant ES cells were thawed and a confluent 3.5 cm well of ES cells was prepared as mentioned earlier. At least 24 hours before blastocyst injection, the cells were split 1:2 and cultured until 70-80% confluency.

Day of blastocyst injection

The medium was changed 3-4 h before trypsinization. The cells were washed twice with PBS, 0.5 ml of trypsin was added, and the cells were then incubated at 37°C for 5 min. 0.5 ml of ES cell medium (with 15-20% serum) was added followed by one volume of serum (0.5 ml). The cells were pipetted 5-10 times to break up clumps and checked under the microscope to make sure there were mostly single cells. Finally, the cells were transferred to screw-cap tubes and placed on ice for blastocyst injection.

5.2.16 Generation of chimeric mice by microinjection

Microinjection was performed at the Max Planck Institute of Experimental Medicine by Mrs. Monika Schindler and animal breeding was performed by the Animal Facility of the Max Planck Institute of Biophysical Chemistry.

5.2.17 Preparation of genomic DNA from the tails of chimeric mice and genotyping by PCR

A piece of tail (about 0.5 cm) was cut off from a 8-week-old mouse (ready for crossing) and put in a 1.5 ml Eppendorf tube. 500 µl of lysis buffer (prepared as in section 5.2.12) was added to cover the tail fragment and incubated at 55°C overnight. The sample was then centrifuged for 5 min at maximum speed and the supernatant was carefully transferred to a new Eppendorf tube. The genomic DNA was purified by extraction with an equal volume phenol/ chloroform, spun down at maximum speed for 5 min and the upper solution was carefully transferred to a tube containing 500 µl of isopropanol, mixed gently and allowed to stand at room temperature for a few minutes. After centrifugation for 5 minutes, the supernatant was discarded. The DNA pellet was washed with 500 µl of 70% ethanol, and centrifuged again for 5 minutes. The supernatant was aspirated and the samples air dried. The DNA pellet was dissolved in 100 µl ddH₂O, and incubated at 37°C overnight. The concentration of genomic DNA was determined and PCR performed as in the section ES cell screening (section 5.2.14).

5.2.18 Crossing of the chimeric mice with wild type mice to generate heterozygotes

Male chimera were used to mate with female wildtype mice to produce heterozygotes. To this end, each male chimera was mated with two female wildtype mice.

5.2.19 Preparation of genomic DNA from the tails of possible heterozygotic mice and verification by PCR

Genomic DNA from the tails of offspring was extracted and genotyped by PCR as in section 5.2.17.

5.2.20 Recrossing of male and female heterozygotes to produce homozygotes

After genotyping, the heterozygotes were recrossed to produce homozygotes. Each male heterozygote was mated with two female heterozygotes.

5.2.21 Genotyping of chimera, heterozygotes and homozygotes by Southern blotting

Southern blotting was performed using an internal probe (cDNA fragment including exon 3 and half of exon 2 amplified with primers P294/P279) and an external probe as shown in the result section (fig 6.1.19).

DNA digestion and transfer:

15-20 µg of genomic DNA extracted from mice tails was digested by SpeI or HindIII, in order to digest completely, a 4-5 fold high concentration (compared with normal digestion) of SpeI or HindIII was used (Roche, 40 u/µl), and the reaction was incubated at 37°C overnight. The samples were run on 0.9% agarose gel at 50 V overnight to separate the digested products. The DNA on the gel was then transferred to QIABRANE membrane (QIAGEN) by the classical Southern transfer method.

The gel was stained in ethidium bromide solution and photographed. The gel was soaked in 0.25 mM HCl for 7 min with agitation to cleave high molecular weight DNA, and then soaked for 30 to 60 min in 1.5 M NaCl-0.5 M NaOH with shaking to denature DNA. The gel was further soaked for 30 to 60 min in 1.5 M NaCl-1.0 M Tris/HCl, pH 8.0 with shaking to neutralize the NaOH. A piece of Whatmann 3 MM paper was placed on a flat support in a dish (reservoir) which was then filled with 10 × SSC. Air bubbles were smoothed out with a glass rod. The gel was placed on damp 3 MM Whatmann paper, inverted. The membrane (Nitrocellulose or Nylon) was cut 1-2 mm wider and longer than the gel. The membrane was wet first with water, then with 2 × SSC for 2-3 min, and then placed on the gel. 2 pieces of 3 MM Whatmann paper in the same size of the gel, were wet in 2 × SSC and placed on top of the membrane. The gel was surrounded with Saran wrap to avoid short-circuiting. A stack of paper towels, 5-8 cm high, was placed on top of the Whatmann paper, followed by a glass plate and a 500 g weight. DNA was transferred for 12-24 hours, changing towels periodically. Towels were removed, the gel and membrane were inverted and gel slots were marked with ball point pen or pencil. The gel was stained in EB to be sure that DNA was transferred. Then the membrane was soaked in 6 × SSC for 5 min, air dried at RT and baked for 2 hours at 80°C (or cross-link by placing on Saran wrap with face down, on UV light for 2 to 5 min).

Pre-hybridization:

Sonicated non-homologous salmon sperm DNA (10 mg/ml) was denatured by heating at 95-100°C for 5min, chilled on ice and added to the solution prepared as below to a final concentration 0.02 ng/ml (50 µl). 25 ml of prehybridization solution was transferred to a hybridization bottle and the membrane was put into the bottle. Pre-hybridization was performed in a hybridization oven at 65°C overnight.

Preparation of pre-hybridization solution as below

Solution	Volume
20×SSC	7.5 ml
50×Denhardt's solution	2.5 ml
5% SDS	2.5 ml
ddH ₂ O	12.5 ml
total	25 ml (for one membrane in one hybridization bottle)

Hybridization:

DNA probe was labelled with the Random Primers DNA Labeling System (Invitrogen) according to the manufacturer's instructions.

25 ng of DNA was dissolved in 5-20 µl distilled water in a microcentrifuge tube and denatured by heating for 5 min at 95-100°C, then immediately cooled on ice. 2 µl dATP solution (0.5 mM dATP), 2 µl dGTP solution (0.5 mM dGTP), 2 µl dTTP solution (0.5 mM dTTP), 15 µl Random Primers Buffer Mixture, and 5 µl of [α -³²P]dCTP (approximately 50 µCi) were added, and then distilled water to a total volume of 49 µl and mixed briefly. 1 µl (3 u/µl) Klenow Fragment was added to the above tube and mixed gently but thoroughly, and centrifuged briefly. The samples were then incubated at 25°C for 1 h and the reaction was stopped with 5 µl Stop Buffer (0.5 M EDTA, pH 8.0). The labelled DNA probe was denatured at 95-100°C for 5 min, then placed on ice, spun down briefly and transferred to the pre-hybridization solution. Hybridization was carried out overnight at 65°C. The hybridised membrane was then washed as follows: 50 ml 2×SSC, 65°C, 15 min; 50 ml 2×SSC-0.1% SDS, 65°C, 30 min; 50 ml 0.1 × SSC, 65°C, 10 min. To exposure, the washed membrane was wrapped in Saran wrap and transferred into a BAS 2040 cassette (FUJI), face upside, covered with BAS-MP2040 film (FUJI) and exposed for 4-5 days. Signals were scanned using a FUJIFILM BAS-2050 device (FUJI) and analyzed using the Basread program (FUJI).

5.2.22 Immunohistochemistry on paraffin sections

The E12 embryos were embedded in wax and cut into 6-8 µm serial sections. The sections were dewaxed in the following steps: 2 × 3-5 minutes in Histochoice clearing agent

(Sigma); 1 × 2 minutes in 100 % ethanol; 1 × 2 minutes in each of 90 %, 70% and 50 % ethanol; 1 × 5 minutes in PBS.

First Antibody:

Slides were put into immuno cassettes and washed with PBS three times. Then the sections were blocked in blocking solution (10% FCS in PBS) for 1 hour at room temperature. Liquid was removed from each section by aspiration or with tissue paper. Sections were framed with a "wax-pencil", covered with the primary antibody diluted in blocking solution (1:50) and incubated overnight at 4°C.

Second antibody:

The first antibody solution was removed from the edge of each slide. The sections were washed, with PBST (0.1 % Tween-20 in PBS) three times, 10 minutes for each with gentle rocking. Then the sections were covered with Cy3-coupled secondary antibody (diluted 1:100 in 10% FCS), and incubated for 0.5-1 hour at RT.

DAPI counter-staining

Sections were washed with PBST three times for 5-10 minutes. Then 1 × bisbenzimidazole was added to the sections and the sections were allowed to stand at RT for 1 min for nucleus staining, and then washed with PBST twice, 5 minutes each. The slides were covered with DAKO fluorescent mounting medium (DAKO Corporation) and stored in the dark space at 4°C. The sections were observed through a microscope (Olympus BS60) using suitable channels (red channel for Cys-coupled antibody staining and blue channel for DAPI counter-staining).

100× Bisbenzimidazole: 0.01% (w/v) in PBS

5.3 Methods for structural and functional analysis of Wind (This work was performed in collaboration with Dr. Qingjun Ma, Dr. Isabel UsÓn and Prof. G.M. Sheldrick at the Department of Structural Biology, University of Göttingen)

5.3.1 Construction of His-Wind and Wind-His expression vectors

Drosophila windbeutel cDNA encoding the full-length Wind was amplified by PCR from a lambda-ZAP library (kind gift of M. Takamori, Max Planck Institute of Biophysical Chemistry). For bacterial expression, a BamHI/SacI PCR product thereof, encoding mature Wind, was ligated into pQE30 (Qiagen), generating an N-terminal extension including a His₆-Tag (MRGSHHHHHHGS). This construct is referred to as His-Wind. A further construct harbouring C-terminally tagged Wind (Wind-His) was prepared by ligating PCR products into pQE60 (Qiagen) using NcoI/BglII sites.

5.3.2 Expression of His-Wind and Wind-His in XL1-Blue

Vectors pQE30-Wind and pQE60-Wind were transformed into XL1-Blue cells. A single clone was picked into 5 ml LB medium with ampicillin and cultured at 37°C, 200 rpm, overnight. 5 ml of an overnight culture was transferred to a flask containing 500 ml LB medium with ampicillin, further cultured until OD₆₀₀≈0.7 (about 4 hours). IPTG was added to 1 mM final concentration to induce protein expression for 2-4 hours. The cells were spun down at 7000 g for 20 minutes. The pellet was resuspended with 40 ml of PBS containing 2 mM Pefabloc (proteinase inhibitor; BioMol), frozen in liquid nitrogen and stored at -20°C for later protein purification.

5.3.3 Purification of His-Wind and Wind-His

Buffers used:

Sonication buffer: 20 mM Tris-Cl (pH 8.0) (RT) 150 mM NaCl	Buffer I: 0.1% Triton X-100 in sonication buffer
Buffer II: 0.35 M NaCl in sonication buffer	Buffer III: 50 mM Pipes (pH 5.7) 100 mM NaCl
Buffer IV: 0.01% (v/v) Triton X-100 8 mM Imidazole (fresh) in sonication buffer.	

Elution buffer: 20 mM Tris-Cl (pH 8.0) (RT)

50 mM NaCl

100 mM Imidazol

0.05% (v/v) Triton-X-100

Dialysis buffer: 10 mM Hepes (pH7.5)

50 mM NaCl

0.01% 2-mercaptoethanol

Purification by Ni-NTA agarose

The cells were thawed in a water bath of RT. 2 ml of 2 M Tris-Cl (pH 8.0) was added to 40 ml of cell resuspension to adjust pH to nearly 8.0. Hen egg white lysozyme was added to a final concentration of 0.75 mg/ml. The cell suspension was incubated at room temperature for 30 minutes to disrupt the cell wall. Then the cells were sonicated with four bursts of 30 seconds, each with 30 second intervals, followed by centrifugation at 12000 g for 20 minutes. Ni-NTA agarose (2 ml solution for 40 ml of cells) was prewashed with buffer I three times, then the supernatant was added to the agarose and incubated for 30-60 minutes at RT with end-over-end rotation. The samples were spun at 7000 g for 10 min and the resin was washed twice with 40 ml of buffer I. The resin was then transferred to a 12 ml MoBiTec gravity-flow column and washed with buffers in the following order: 3 × bed volumes of buffer I, 4 × bed volumes of buffer II; 3 × bed volumes of buffer I, 4 × bed volumes of buffer IV. Proteins were eluted with 5-10 bed volumes of elution buffer and fractions of the elution were collected. The protein concentration of each fraction was determined and the fractions with peak concentration were collected.

5.3.4 Dialysis and concentration

The purified recombinant protein was transferred into a suitable dialysis bag (cut off 8,000-12,000 Dalton) and dialysed thoroughly against the dialysis buffer at 4°C. After dialysis, the protein was concentrated to a final concentration of 24 mg/ml, frozen in liquid nitrogen and stored at -20°C unless used immediately.

Dialysis buffer: 10 mM HEPES pH 7.5

50 mM NaCl

0.01% 2-mercaptoethanol

5.3.5 Immunisation of rabbits to generate antibodies against His-Wind

(Sample preparation and injection of rabbits were performed by the staff of the Animal Facility of Max-Planck-Institute of Biophysical Chemistry)

For the initial injection, 1 mg of recombinant protein, His-Wind, was used. Boosts, given after eight weeks and thereafter as required, contained 0.5 mg of recombinant protein, but were otherwise similarly prepared.

5.3.6 IgG purification using Protein A Sepharose

IgG fractions from serum were purified over Protein A Sepharose columns as follows: 1 ml of serum was spun for 30 min at 14,000 g in a tabletop centrifuge. The pH of the supernatant was set to 8.1 with 1 M Tris/HCl pH 9.0, and the sample was then applied to a column holding 1 ml of Protein A Sepharose pre-washed with 10 ml Pi buffer. Binding of IgG proteins was allowed to proceed overnight by slowly (about 0.25 ml/min) circulating the serum repeatedly over the column with the aid of a peristaltic pump. The column was then washed with buffer Pi until all unbound proteins had been washed out (measured at A_{280}). Bound IgG was eluted in one step with Elution buffer. Eluted proteins were collected in 1.5 ml tubes already containing sufficient neutralization buffer (about 300 μ l) for immediate neutralization. The fractions were further aliquoted and frozen or used for further purification as described below.

Pi buffer: 0.1 M sodium phosphate, pH 8.0
0.01% NaN₃

Neutralisation buffer: 1 M Tris/HCl, pH 9.0

Elution buffer: 0.2 M glycine, pH 3.5

5.3.7 Removal from anti-Wind antiserum of antibodies against *E.coli* protein using CNBr activated Sepharose 4B

Binding buffer: 100 mM NaBH₃, pH 8.0
1 M NaCl

This method was used for larger amount of antibody solutions. Cells from a 1 L culture of the appropriate *E. coli* strain (XL1-Blue or BL21), grown to stationary phase, were

collected by centrifugation, resuspended in 100 ml binding buffer and treated for 20 min at RT with 2 mg/ml lysozyme. For further solubilisation and to degrade DNA, Triton X-100 and DNase were added to 0.2% (w/v) and 10 µg/ml respectively, and further incubated at 4°C for 1 h. The cleared suspension was centrifuged at 8,000 g and 4°C for 20 min, and the pH of the supernatant was adjusted to 9.0 with NaOH. The solution was chilled on ice, and proteins were then coupled to CNBr-activated Sepharose 4B according to the manufacturers recommendations. The bound material was washed with 0.2 M glycine, pH 8.0, then with TBS containing 0.2% sodium azide, and stored at 4°C until required.

1 ml settled volume of this material was used per mg IgG prepared as in section 3.3.7. The IgG preparation was added to the coupled material and incubated with agitation for 8-12 h at RT. Purified antibodies were recovered by draining the slurry over a suitable column (e.g. a 5 ml MoBiTec column, MoBiTec). The column was washed with TBS until the A280 had dropped almost to zero (or monitored at A598 if Bradford method was used). 0.2 column volumes were collected. The antibodies were pooled, aliquoted and stored at -20°C.

5.3.8 Insulin reduction assay

Reduction of the interchain disulfide bonds in insulin leads to a precipitation of the free B-chain, while the A-chain remains in solution. The resulting turbidity can be followed spectrophotometrically at 600 nm. Here, the assay was used to test a possible reductase activity of purified, recombinant *Drosophila* proteins His-Wind and Wind-His. The PDI family CaBP2 was used as a positive control. Before the experiment, 2-ME was eliminated from the protein solutions by filtration over G-25 column (Pharmacia). The proteins were eluted from the column in PBS buffer, the same amount of PBS buffer was used as a negative control.

Potassium phosphate buffer: 0.1 M potassium phosphate; pH 6.5
2 mM EDTA

Tris/HCl buffer: 0.05 M Tris/HCl; pH 8.0

50 mg of insulin was dissolved in 4 ml Tris/HCl buffer, the pH adjusted to 2-3 with 1.0 M HCl and rapidly titrated back to 8 with 0.1 M NaOH. The solution, adjusted to 5 ml, was stored at -20° C. Insulin from the stock solution was diluted to 1 mg/ml with potassium phosphate buffer, His-Wind/Wind-His was added to a final concentration of 0.28 µM in 600

μl and the reaction started with 4 mM GSH (final concentration). The reaction was followed continuously at 600 nm in a Shimadzu UV 160A spectrophotometer.

5.3.9 Cross-linking of His-Wind/Wind-His

Cross-linking by D.M. Ferrari was performed to investigate the oligomerisation properties of recombinant His-Wind. The cross-linker used in this work was glutaraldehyde. About 71 μg of His–wind and Wind-His were diluted into 250 μl of 50 mM sodium borate (pH 7.4) and added to reaction tubes containing different amounts of NaCl (0, 50, 150 and 500 mM). Water was then added to a total volume of 500 μl. The reaction mix was incubated for 20-30 min at RT. 20 μl of 25% glutaraldehyde (final concentration 1%) was added to all samples. The samples were well mixed and incubated for 2 min at RT. The reactions were quenched with 12.5 μl of stop solution and incubated for 20 min at RT. Proteins were precipitated by adding 30% (w/v) TCA to a final concentration of about 8% and the pellets were washed with cold acetone. The protein pellets were dissolved in Laemmli loading buffer and samples were run on 10%SDS-PAGE gel.

Solutions: 50 mM sodium borate pH 7.4

Stop solution: 2 M NaBH₄
0.1 M NaOH

5.3.10 Structure determination of Wind

5.3.10.1 Crystallization

The initial crystals appeared in No.24 of Hampton Screen 2 (Hampton). After optimizing protein concentration, precipitant, buffer, temperature, pH and additives, the best crystals usable for diffraction were grown with the hanging-drop vapor diffusion method. The drops consisted of 6ul [5.8mg/ml Wind in 5mM Hepes, pH 7.5, 25mM NaCl, 0.0025% (V/V) BME] and 3ul well solution [0.1M MES, pH6.1,0.1M CsCl, 2mM CaCl₂, 16% PEG300]. 0.1M MES, pH 6.1, 0.1M CsCl (or NaCl), 20% PEG 300, 10% glycerol was used as cryo-solution for data collection at low temperature.

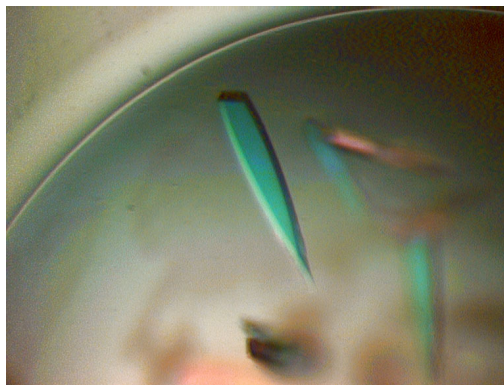


Figure 5.1 Crystal picture of Wind after optimization.

5.3.10.2 Heavy atom derivative

The Hg derivatives were prepared by growing crystals with the well solution of 0.1M MES pH 6.1, 0.08 M NaCl, 2 mM CaCl₂, 16% PEG 300, and soaking them in 0.1 M MES pH 6.1, 0.1 M NaCl, 20% PEG300, 10% glycerol, 0.05 mM HgCl₂ for 2 weeks.

5.3.10.3 X-ray data collection and processing

The native data were collected by Q. Ma on a Mar CCD detector at the synchrotron beamline X11 in Hamburg at the EMBL outstation DESY. The Hg derivative data and another native data set were collected on the Mar345 detector with Cu rotating anode in the Department of Structural Chemistry, University of Göttingen.

5.3.10.4 Data analysis, model building, refinement, and structural analysis were performed in collaboration with Dr. Qingjun Ma, Prof. Dr. Isabel Uson, and Prof. George M. Sheldrick at the Department of Structural Biology, University of Göttingen and is reported elsewhere (Ma *et al.*, 2003. J. Biol. Chem in press).

6. Results

6.1 Studies on ERp28

6.1.1 Gene structure of mouse ERp28

At the time when I started the project, the mouse ERp28 cDNA was unknown and the sequencing of the mouse genome had not been finished. Therefore the first step was to look for the cDNA sequence of mouse ERp28. The cDNA sequence of human ERp28 was used to search the mouse ESTs database and several ESTs with high homology to human ERp28 were found. Through alignment of these ESTs, almost the full length of mouse ERp28 cDNA including the complete coding region was obtained. Below is the cDNA sequence of mouse ERp28:

```
ggtagcttgccctgtctctcccgctcccaccatccggcgtgatgggccgcccgcgc
cgggggtgtctggcgctgcttctctctccccactgctgtccgtgctcctgggcct
gctgctcctctttgctccgcacggcgccagcggcctacacacgaaggggtgcctt
tcccttggaacacagtcactttctacaaggtcattcccaaaagcaagttcgtctt
ggtgaagttcgacacccagtacccctatggagagaaagcaagatgagttcaagcg
tctggctgagaactcagcctccagcagaggagctcttgggtggcagaggtggggat
ctcagactatggcgacaagctgaacatggagctgagtgagaagtacaagctgga
caaggagagctacccagtcttctacctattccgggatggggacttggagaaccc
tgtgctgtacaatggggcagttaagggttgagccatccagcgtggctaaaggg
gcaggggggtctacctgggcatgcctgggtgcctgcctgcatacgatgccctggc
aggcaggttcatcaaggcctccagcatagaggcccgacaggccatcttgaaaca
gggacaggatggcctcctaagtgtgaaggaaacagagaagaagtgggctagcca
atacctgaagatcatggggaagatcctggaccaaggtgaggactttccagcctc
tgagatggcccggatcggttaagctgatagagaacaagatgagtgacagcaagaa
agaagagctacagaagagcttaaacatcctgacggccttccggaagaaggaggc
tgagaaggaggagctgtgaccaagcacaaagggttccccagggtttgccaggggc
aggaaaggagagatctacctgccagctgtgggacccctgtgggtggaaggggca
gtggagcaatgcaggcctgagccagaagtctgtgccccgagtgctgctggacat
tgatgctgctgagatcatacttaggacacccccctgagccagtctgtggagagcc
tgtgtgctcagtcagatggctggtgagatgccccccgcaggagttggtgctat
agaggtagggactactgccaggctccttgatgagtgtaattctgatcaaataaaa
agtctgttttgggt
```

Figure 6.1.1: The cDNA sequence of mouse ERp28 containing three exons. Exon2 is underlined and in italics. The translation start codon atg and the stop codon tga are in bold type.

The second step was to look for the gene structure (exon distribution in the ERp28 locus). For this purpose, the cDNA sequence of mouse ERp28 was used to search the mouse genomic DNA database, the exon/intron junctions were determined by comparing the cDNA sequence with some known genomic fragments containing the locus of ERp28. The gene of mouse ERp28 contains three exons: The first exon contains the translation initiation codon

ATG and a short 5' untranslated region before the ATG; the third exon contains the stop codon TGA and the 3' untranslated region after the TGA. Primers were designed from the first and the second exons (P278/P293) and the second and the third exons (P294/P279), respectively. The distance between the exons (introns) could be worked out according to the size of the PCR products. The picture below shows that the first intron is about 5 kb and the second about 2.5 kb in length.

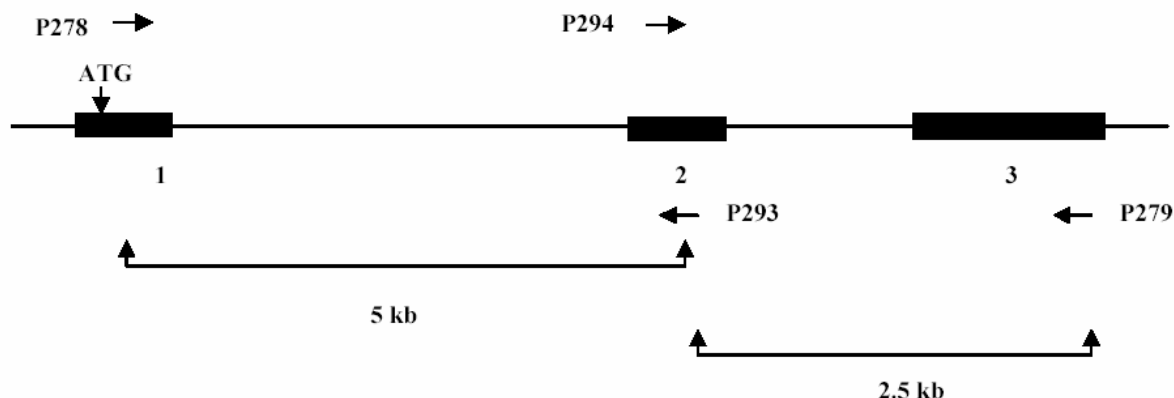


Figure 6.1.2: The gene structure of mouse ERp28. Exons are depicted with filled boxes and the distances between exon1 and exon2, exon2 and exon3 are about 5kb and 2.5kb, respectively.

6.1.2 The amino acid sequence of mouse ERp28 and the similarity with the human protein

The mouse ERp28 cDNA was translated into amino acids and the [PSORT II Prediction](http://psort.nibb.ac.jp/form2.html) program (<http://psort.nibb.ac.jp/form2.html>) was used to predict the signal sequence, while the secondary structure was predicted by the PSIPRED Protein Structure Prediction Server (<http://bioinf.cs.ucl.ac.uk/psipred/>). The results show that the full length protein has 262 amino acids (fig 6.1.3) and contains a signal peptide at the N-terminus (amino acid 1-34, underlined), with a cleavage site between G34 and L35. At the very C-terminus, as in human ERp28, is a KEEL (in italics) ER retrieval signal. The identity between human ERp28 and mouse ERp28 is about 90%. Interestingly, both human and mouse ERp28 carry a second internal KEEL motif (in bold). The function of this second motif remains to be established.

MAAAAGVSGAASLSPLLSVLLGLLLLFAPHGGSGLHTKGALPLDTVTFYKVI PK
 SKFVLVKFDTQYPYGEKQDEFKRLAENSASSEELLVAEVGISDYGDKLNMEISE
 KYKLDKESYPVFYLF RDGDLENPVLYNGAVKVGAIQRWLKGQGVYLGMPGCLPA
 YDALAGEFIKASSIEARQAILKQGQDGLLSVKETEKK WASQY LKIMGKILDQGE
 DFPASEMARIGKLIENKMSDSK**KEEL**QKSLNILTAFRKKEAE**KEEL**

Figure 6.1.3: The amino acid sequence of mouse ERp28 protein. The signal sequence is underlined. The C-terminal ER retrieval signal KEEL is in bold italics and the internal KEEL in bold type.

A secondary structure analysis (data not shown) predicts that mouse ERp28 protein, like human ERp28, contains two domains; an N-terminal thioredoxin-like domain (hereafter referred to as the b domain) and a C-terminal D domain with five α -helices which is unique to the PDI-D subfamily.

6.1.3 RT-PCR and the possible alternative splicing form of mouse ERp28

The first step in knocking out a gene is to build a knock out vector. In the vector, two arms (left and right) allow specific homologous recombination with the mouse genome. To obtain the two arms, a mouse genomic DNA library (strain S129) was screened with a [α - 32 P] labelled cDNA probe of mouse ERp28 prepared by RT-PCR. RT-PCR was performed using total RNA extracted from brain tissue of E12.5 mouse embryos as template, and P278/P294 as primers. Unexpectedly, the reaction produced two bands with distinct sizes running very close to each other on agarose gels (fig 6.1.4 A). The same results were also observed for the products using total RNA extracted from cultured F9 cells (fig 6.1.4 B). After TA cloning and sequencing, the upper band (arrows in fig 6.1.4) was proven to be the cDNA sequence of mouse ERp28; the lower band (arrowheads in fig 6.1.4) was shown to contain the first and third exon, but not the second exon of full length mouse ERp28 cDNA (AK009881). Searching the mouse EST database using the sequence of the lower band showed that several ESTs do exist which can well match the sequence of the lower band (data not shown). This indicates that the short form is not an artificial product and might be an alternative splicing form of mouse ERp28.

An alignment result between one of the short form EST (AK013303) and the full length cDNA of mouse ERp28 (AK009881) is shown in figure 6.1.5 [the upper line: alternative splicing form; the lower line: mouse ERp28 cDNA]. Only the sequences before the stop codon TGA are shown although the sequences after TGA also match very well. From the alignment, one can clearly see that the second exon containing a splicing signal (GT-AG, underlined) is spliced out from the cDNA of mouse ERp28.

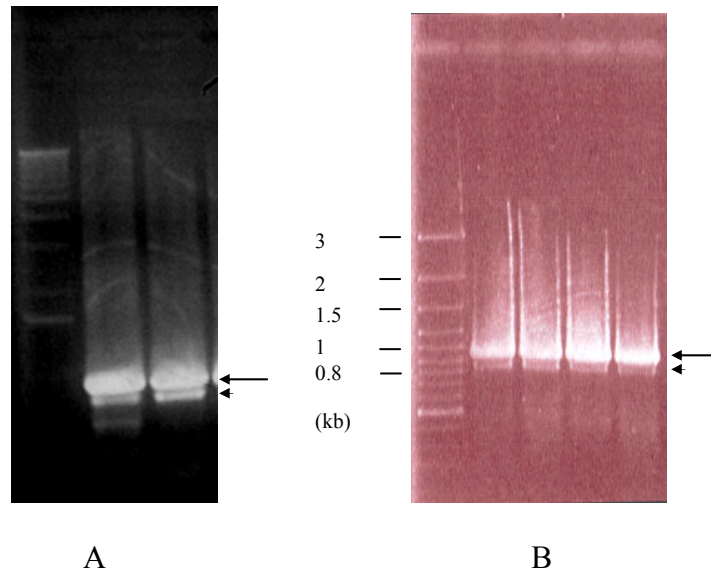


Figure 6.1.4: RT-PCR.

A: RT-PCR from total RNA extracted from E12.5 brain; B: RT-PCR from total RNA extracted from F9 cells. The different lanes in both pictures are RT-PCR products under different conditions. The arrows and arrowheads represent the RT-PCR products of mouse ERp28 and the alternative splice form, respectively. The size of 100 bp marker (MBI) is shown on the left side of fig 6.1.4 B.

```

TG TAGCTTGCCTGTCTCTCCCGCTCCCAACATCCGGCGTGATGGCCGCCGCCGCCGGGGTGTCTGGCGCTG
GTAGCTTGCCTGTCTCTCCCGCTCCCAACATCCGGCGTGATGGCCGCCGCCGCCGGGGTGTCTGGCGCTG

CTTCTCTCTCCCACTGCTGTCCGTGCTCCTGGGCCTGCTGCTCCTCTTTGCTCCGCACGGCGGCAGCGGC
CTTCTCTCTCCCACTGCTGTCCGTGCTCCTGGGCCTGCTGCTCCTCTTTGCTCCGCACGGCGGCAGCGGC

CTACACACGAAGGGTGCCCTTCCCTTGGACACAGTCACTTTCTACAAG-----
CTACACACGAAGGGTGCCCTTCCCTTGGACACAGTCACTTTCTACAAGGTCATTCCCAAAGCAAGTTCGT

-----
CTTGGTGAAGTTCGACACCCAGTACCCCTATGGAGAG AAGCAAGATGAGTTCAAGCGTCTGGCTGAGAAC

----- ACTATGGCGACAAGCTGAACATGGA
TCAGCCTCCAGCGAGGAGCTCTTGGTGGCAGAGGTGGGGATCTCAGACTATGGCGACAAGCTGAACATGGA

GCTGAGTGAGAAGTACAAGCTGGACAAGGAGAGCTACCCAGTCTTCTACCTATTCCGGGATGGGGACTTGG
GCTGAGTGAGAAGTACAAGCTGGACAAGGAGAGCTACCCAGTCTTCTACCTATTCCGGGATGGGGACTTGG

AGAACCCTGTGCTGTACAA TGGGGCAGTTAAGGTTGGAGCCATCCAGCGCTGGCTAAAGGGGCAGGGGGT
AGAACCCTGTGCTGTACAA TGGGGCAGTTAAGGTTGGAGCCATCCAGCGCTGGCTAAAGGGGCAGGGGGT

CTACCTGGGCATGCCTGGGTGCCTGCCTGCATACGATGCCCTGGCAGGCGAGTTCATCAAGGCCTCCAGCA
CTACCTGGGCATGCCTGGGTGCCTGCCTGCATACGATGCCCTGGCAGGCGAGTTCATCAAGGCCTCCAGCA

TAGAGCCCCGACAGGCCATCTTGAAACAGGGACAGGATGGCCTCCTAAGTGTGAAGGAAACAGAGAAGAAG
TAGAGCCCCGACAGGCCATCTTGAAACAGGGACAGGATGGCCTCCTAAGTGTGAAGGAAACAGAGAAGAAG

TGGGCTAGCCAATACCTGAAGATCATGGGGAAGATCCTGGACCAAGGTGAGGACTTTCAGCCTCTGAGAT
TGGGCTAGCCAATACCTGAAGATCATGGGGAAGATCCTGGACCAAGGTGAGGACTTTCAGCCTCTGAGAT

GGCCCGGATCGGTAAGCTGATAGAGAACAAAGATGAGTGACAGCAAGAAAGAAGAGCTACAGAAGAGCTTAA
GGCCCGGATCGGTAAGCTGATAGAGAACAAAGATGAGTGACAGCAAGAAAGAAGAGCTACAGAAGAGCTTAA

ACATCCTGACGGCCTTCCGGAAGAAGGAGGCTGAGAAGGAGGAGCTGTGA
ACATCCTGACGGCCTTCCGGAAGAAGGAGGCTGAGAAGGAGGAGCTGTGA

```

Figure 6.1.5: Alignment of the short form EST with cDNA of mouse ERp28. ATG and TGA are shown in bold type. The splicing signal GT-AG is underlined.

Interestingly, when I used the sequence of the short form of mouse ERp28 to search the human EST database, I found high similarity with several human ESTs. Analysis of these ESTs (for example: BG913429) showed that the second exon was also missing (data not shown). This result further confirmed the alternative splice form of ERp28 and indicates that it exists not only in the mouse but also in human, and probably in other species.

However, when the cDNA sequence of the alternatively spliced form was translated into amino acids, it turned out that this form was not likely to be translated into a protein if the same initiator ATG as in ERp28 was used due to several nearby translation stop codons. The translation result of the alternative splice form is shown in fig 6.1.6. Clearly, the lack of exon2 causes a reading frame shift, thus it cannot be translated into a protein in the first (used in ERp28) and second frames. However, according to the third frame it may form a protein which has the same amino acid sequence as the C-terminal portion of ERp28 if a different initiator ATG is used.

5'3'	Frame	1
F H P G G G R S I S L W R W P L A F Met A A A G R T A L P V T Stop P P T R S A P P V		
A C L S L P L P P S G V Met A A A A G V S G A A S L S P L L S V L L G L L L L F A P		
<u>H G G S G L H T K G A L P L D T V T F Y K</u> T Met A T S Stop T W S Stop V R S T S W T		
R R A T Q S S T Y S G Met G T W R T L C C T Met G Q L R L E P S S A G Stop R G R G S		
T W A C L G A C L H T Met P W Q A S S S R P P A Stop R P D R P S Stop N R D R Met A		
S Stop V Stop R K Q R R S G L A N T Stop R S W G R S W T K V R T F Q P L R W P G S V		
S Stop Stop R T R Stop V T A R K K S Y R R A Stop T S Stop R P S G R R R L R R R S C		
D Q A Q G F P R V C Q G Q E R R D L P A S C G T P V G G R G S G A Met Q A Stop A R S		
L C P E C A W T L Met L L R S Y L G H P L S Q S V E S L C A Q S R W L V R C P P Q E		
L C Y R G R D Y C Q V L D E C N S D Q I K S L F W		
5'3'	Frame	2
S T P E A A G A S R S G G G R L R S C A P P A G R P S P Stop R D R P L G A L R L		
Stop L A C L S R S H H P A Stop W P P P P G C L A L L L S P H C C P C S W A C C S S		
L L R T A A A Y T R R V P F P W T Q S L S T R L W R Q A E H G A E Stop E V Q A G Q		
G E L P S L L P I P G W G L G E P C A V Q W G S Stop G W S H P A L A K G A G G L P G		
H A W V P A C I R C P G R R V H Q G L Q H R G P T G H L E T G T G W P P K C E G N R		
E E V G Stop P I P E D H G E D P G P R Stop G L S S L Stop D G P D R Stop A D R E Q		
D E Stop Q Q E R R A T E E L K H P D G L P E E G G Stop E G G A V T K H K G S P G F		
A R G R K G E I Y L P A V G P L W V E G A V E Q C R P E P E V C A P S V P G H Stop C		
C Stop D H T Stop D T P Stop A S L W R A C V L S Q D G W Stop D A P R R S C A I E V		
G T T A R S L Met S V I L I K Stop K V C F G		
5'3'	Frame	3
P P R R R P E H L A L A V A A C V H A R R R Q D G P P R D V T A H S E R S A C S L P		
V S P A P T I R R D G R R R R G V W R C F S L P T A V R A P G P A A P L C S A R R Q		
R P T H E G C P S L G H S H F L Q D Y G D K L N Met E L S E K Y K L D K E S Y P V F		
<u>Y L F R D G D L E N P V L Y N G A V K V G A I Q R W L K G Q G V Y L G</u> Met P G C L P		
<u>A Y D A L A G E F I K A S S I E A R Q A I L K Q G Q D G L L S V K E T E K K W A S Q</u>		
<u>Y L K I</u> Met G K I L D Q G E D F P A S E Met A R I G K L I E N K Met S D S K K E E L		
<u>Q K S L N I L T A F R K K E A E K E E L</u> Stop P S T R V P Q G L P G A G K E R S T C Q		
L W D P C G W K G Q W S N A G L S Q K S V P R V C L D I D A A E I I L R T P P E P V		
C G E P V C S V K Met A G E Met P P A G V V L Stop R Stop G L L P G P Stop Stop V		
Stop F Stop S N K K S V L A		

Figure 6.1.6: Translation of the alternatively spliced form of mouse ERp28. The first methionine (the translation start point), which is used in ERp28 is in bold type and underlined as shown in the first translation frame. The first part of amino acids in mouse ERp28 is underlined in the first frame. In the third translation frame, the sequence that may be translated from the alternative splice form is underlined and the methionines which are possibly used as translation start point are in bold italics

6.1.4 Mouse genomic DNA library screening

Three positive colonies were obtained from the first round of screening. They were further confirmed in a second and third round of screening. PCR results using different primers proved that two of them (No.2 and No.3) contained all of the three exons and introns, while the third (No.1) might not because no product could be obtained with primers p323 and p273.

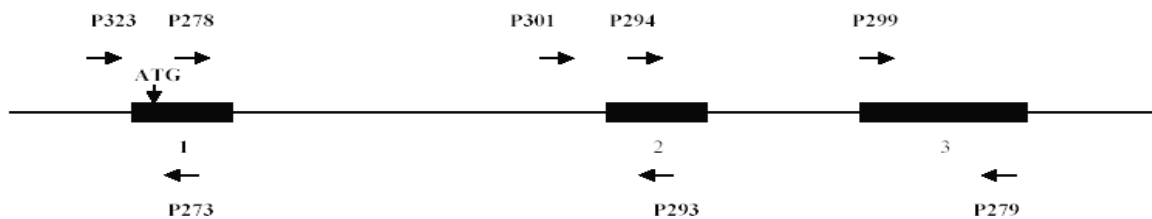


Figure 6.1.7: primers used for screening. Exons are shown as filled boxes.

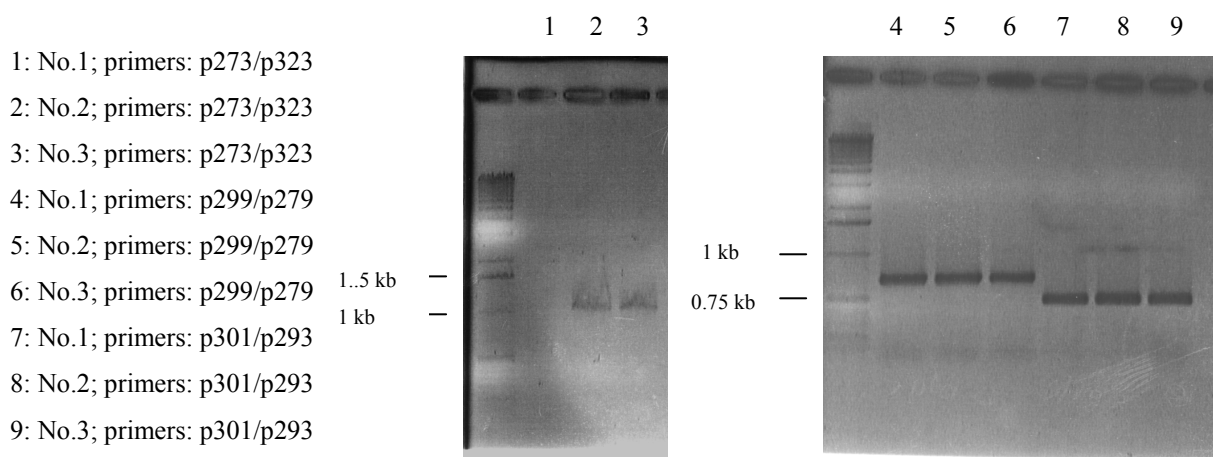


Figure 6.1.8: PCR results from the three positive clones to check whether they contain the whole locus of mouse ERp28. PCR products could be obtained from all the three positive clone with primer pairs p301/p293 and p299/p279 (lane 4-9). However, with primer pair p273/p323, the expected products could be obtained from No.2 (lane 2) and No.3 (lane 3), in contrast, no product (lane 1) could be amplified from the first positive clone (No.1), indicating that No.1 might not contain the region around exon 1. The representative size (kb) of the 1 kb marker (MBI) is marked on the left side of the pictures.

6.1.5 Map of restriction sites in the mouse ERp28 locus

The ERp28 locus was isolated from the positive Lambda FIXII phage DNA (No.2) by digestion with NotI and EcoRI (fig 6.1.9). The resulting three fragments (6 kb, 5 kb, and 4 kb) from the ERp28 locus were cloned into the plasmid pBluescript KS II. Then different

restriction enzymes were used to digest these plasmids. The digestion results were combined to reveal the restriction map of the ERp28 locus. The map is shown in fig 6.1.10. With the map at hand, preparation of the knockout vector could be started. The restriction enzyme digestion and sequencing results showed that the 4 kb fragment contained the upstream portion of the first exon (the vector was named pBS-4kb); the 6 kb fragment included the second and the third exon (pBS-6kb); the 5 kb fragment was in between (pBS-5kb).

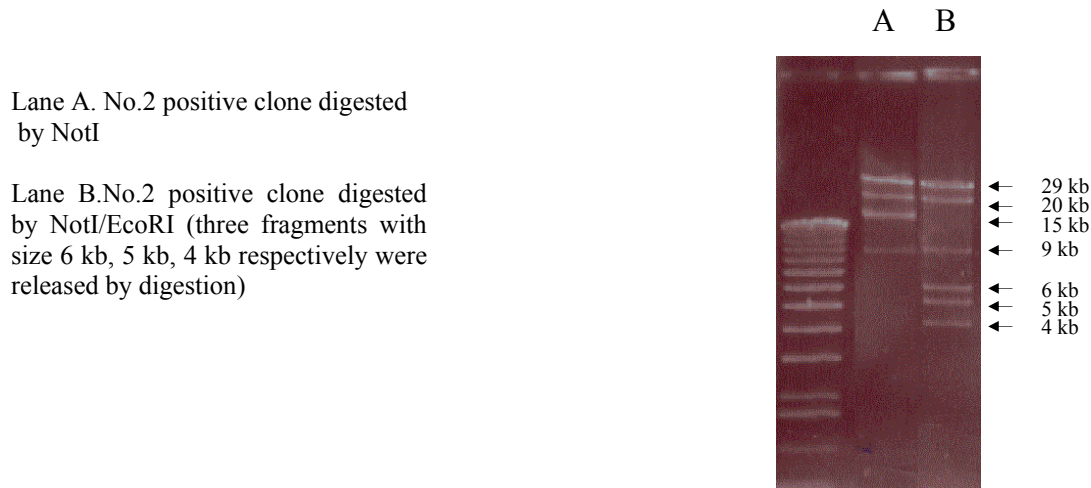


Figure 6.1.9: Isolated and digested ERp28 locus by Not I and EcoR I. Lane A represents a NotI digestion of the positive Lambda FIXII phage DNA. The two upper bands (29 and 20 kb) and lower band (9 kb) show NotI fragments from the vector, whereas the third band (15 kb) from the top contains the ERp28 genomic DNA insert which is further cleaved by EcoRI (lane B) to generate 6, 5 and 4 kb fragments.

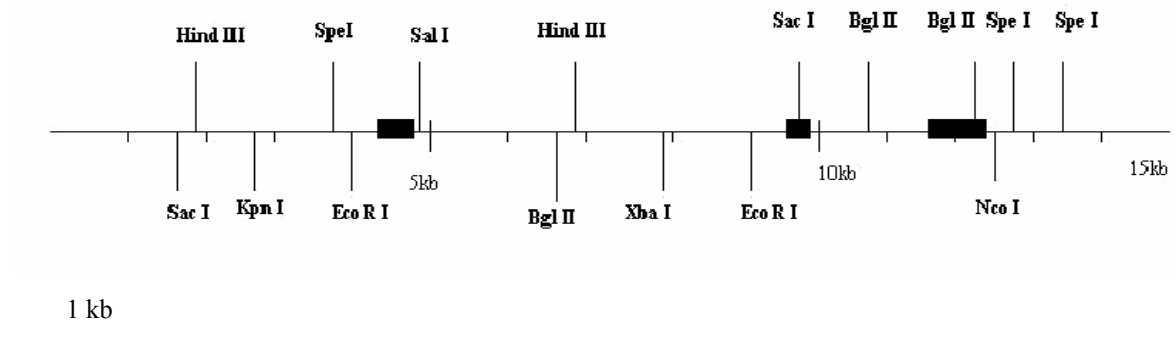


Figure 6.1.10: Restriction map of mouse ERp28 locus. Filled boxes represent the three exons of the mouse ERp28 gene. Restriction enzyme sites are shown as long vertical lines. The scale bar below the picture represent 1 kb in length.

6.1.6 Mouse ERp28 knock out vector composition

The ERp28 knock out vector consists of five parts apart from the backbone: (1) the left arm and (2) the right arm will be used for homologous recombination with the mouse ERp28

locus; (3) the reporter gene LacZ is cloned behind the left arm, and by using the endogenous promoter of ERp28, it will indicate the expression pattern of mouse ERp28. After recombination, LacZ will replace the translation start point ATG and the region behind ATG in the first exon, thus ERp28 can not be translated and its function will be lost (as shown in fig 6.1.11 and fig 6.1.12). (4) the neomycin resistance gene (lox^pneo^{lox}) will be used as a positive selection marker. It is reported that in some cases the neomycin resistance gene may affect the phenotype of knock out mice and can even lead to lethality, therefore two lox^p sites flank the neomycin resistance gene so that if necessary, the neomycin resistance gene can be deleted by the action of a recombinase which can be introduced into positive recombinant ES cells by transfection with a recombinase expression vector. (5) TK (thymidine kinase) is used as a negative selection marker.

The expected recombination is shown in fig 6.1.12.

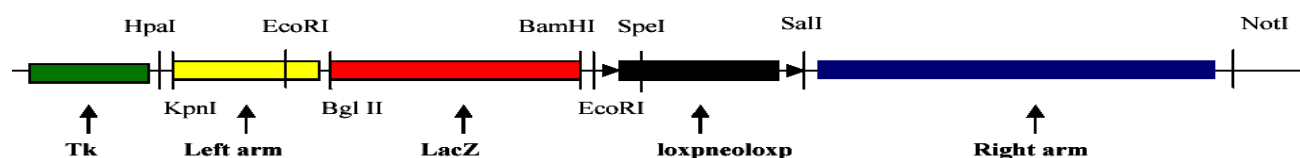


Figure 6.1.11: The mouse ERp28 knock out vector. Tk (thymidine kinase) is shown as green box; Left arm in yellow; Right arm in blue; The red box is lacZ; the loxpneo is in black and the two arrows flanking the black frame are lox sites. The vertical lines represent different restriction enzyme sites.

6.1.7 Steps for making the ERp28 Knock out vector

The knock out vector is relatively complex and many restriction enzyme sites were used. The following steps were made for creating the ERp28 knock out vector (composition as shown in fig 6.1.11):

1. The vector pKO-Scrambler-NTKV-1908 (kindly provided by Dr. X.Zhou) was modified by deletion of the multiple cloning sites between BamHI and SmaI and by insertion of a polylinker from the plasmid pSL-1180 (kindly provided by Dr. X.Zhou) using HpaI and EcoR I sites. This modification provided more suitable restriction enzyme sites for future cloning.
2. The LacZ reporter gene was cloned into the modified pKO-Scrambler-NTKV-1908 using BglII and BamHI sites (pKO-LacZ).
3. The left arm (LA) was prepared and cloned into pKO-LacZ using KpnI and BglII sites. The left arm including the promoter region of mouse ERp28 was prepared in two steps: (1)

the plasmid pBS-4kb (section 6.1.5) was digested by KpnI and EcoRI and the product was cloned into pKO-Scrambler-NTKV-1908; (2) then the region between the EcoRI site (included) and the initiator ATG (not included) was amplified from pBS-5kb (section 6.1.5) by PCR and the PCR product was cloned into the vector above using EcoRI and BglII sites. The insert including both fragments (2 kb) was isolated from this vector by KpnI and BglII and recloned into pKO-LacZ to generate pKO-LacZ-LA.

4. The neomycin resistance gene of pKO-LacZ-LA was removed using SalI and replaced by the loxpneoloxp sequence from the vector pBSloxpnoloxp (kindly provided by Dr. X.Zhou) using BamHI and Sal I sites. This product was termed pKO-LacZ-LA-loxpneoloxp.

5. In the last step, the right arm (RA) was isolated from pBS-6kb (section 6.1.5) (the vector was first digested with EcoRI, blunted with Klenow fragment, then cut with NotI) and then the right arm was cloned into pKO-LacZ-LA-loxpneoloxp using a blunted SalI site and a NotI site (pKO-LacZ-LA-loxpneoloxp-RA). Since the original vector pKO-Scrambler-NTKV-1908 contained the negative selection marker (TK), it did not need to be recloned.

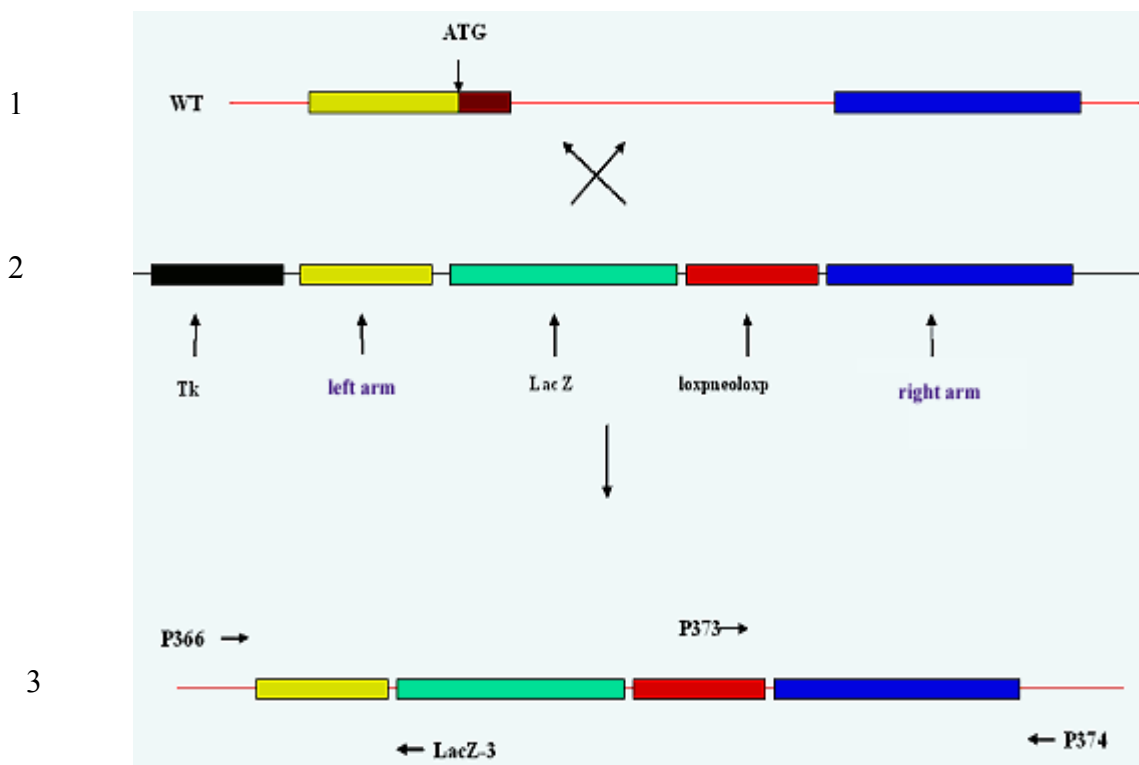


Figure 6.1.12: The wild type locus of mouse ERp28, the knock out vector, the expected recombination and primers for screening the ES cells, genotyping chimera, heterozygotes, and homozygotes. Line1: wild type locus; Line 2: the ERp28 knock out vector; Line 3: the expected recombination result; P366/LacZ-3 and P373/P374 are primers for screening and genotyping, primers P366/LacZ-3, P373/P374 will generate a fragment about 2.3 kb, and 7 kb in size respectively. Similarly colored boxes represent identical elements on all lines.

6.1.8 Selection of transfected ES cells by G418 and Gancyclovir and screening of ES cells by PCR

Transfected ES cells were first selected by simultaneous treatment with G418 and Gancyclovir. 193 surviving ES clones were obtained. Genomic DNA extracted from these ES clones were screened by PCR using two pairs of primers P366/LacZ-3 and P373/P374 (fig 6.1.12). P366 and LacZ-3 were designed from the upstream region of the left arm within the genome and the sequence of LacZ, respectively, and should produce a product of about 2.3 kb. P373/P374 were designed from the sequence of loxpneoloxp and the downstream region of the right arm in the genome, respectively, and should produce a product of about 7 kb. Therefore, primer pairs produce PCR products from positive homologous recombinant ES cells rather than from random recombinant or insertion ES cells. Figure 6.1.13 shows that PCR products could be generated from three ES clones (C27, C61 and C65). These PCR products were then cloned into the TA cloning vector and sequenced. The restriction enzyme digestion and sequencing results further proved that all three ES clones contained true positive recombinations. Genomic DNA from wild type ES cells served as negative control for all of these PCR reactions. The positive recombinant ES cells were confirmed later by Southern blot using external probe (fig.6.1.19 and fig.6.1.21).

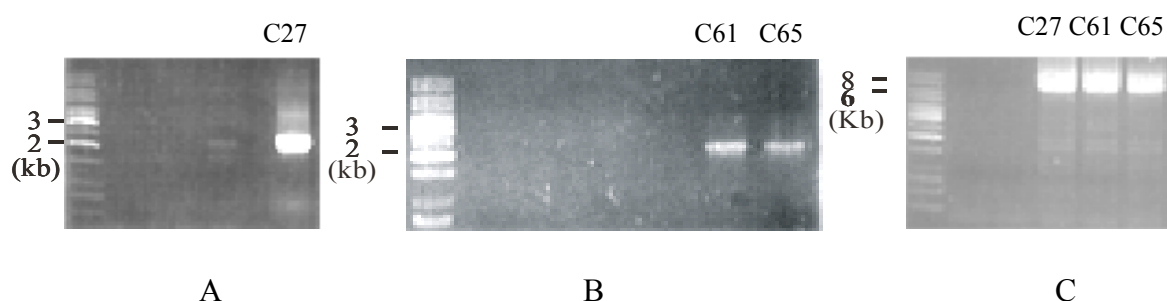


Figure 6.1.13: ES cells screened by PCR. A and B: PCR products (about 2.3 kb) could be amplified from the genomic DNA extracted from three positive recombinant ES cells by primers P366/LacZ-3 (C27, C61 and C65). C: PCR products (about 7 kb) were obtained from the identical ES cells by primers P373/P374. In contrast, no PCR product could be produced from other ES cells as shown in the remaining lanes.

6.1.9 Genotyping of chimeric mice by PCR

To produce chimeric mice, two of the positive ES clones (C27, C61) were microinjected into mouse blastocysts (strain: C57BL6N). In total, 17 chimeric mice (14 male and 3 female) were obtained. The mice were genotyped by PCR using the primers as above

and genomic DNA extracted from the tails of these mice. PCR products could be amplified from all of the chimeric mice and a representative selection is shown in fig 6.1.14. However, the chimeric levels of these mice were quite different, reflected in their coat color. The stronger chimeric mice have a greater chance of producing germline cells which develop from the recombinant ES cells.

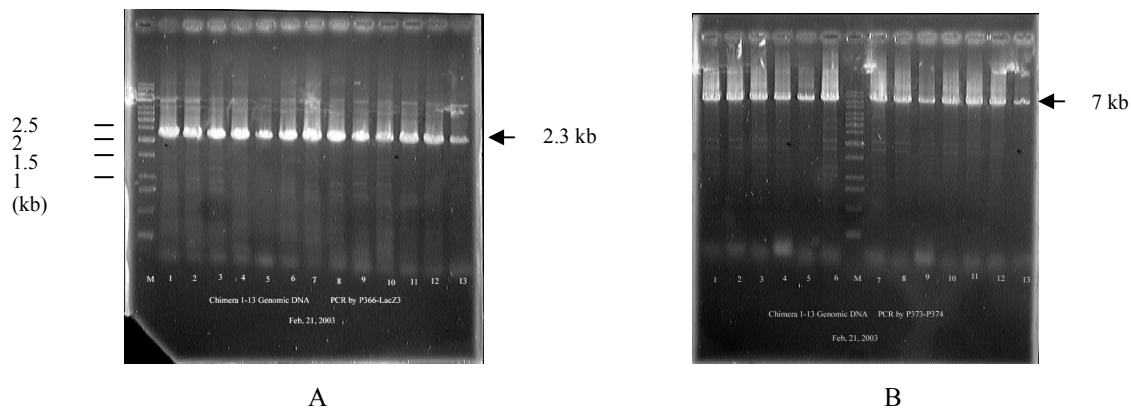


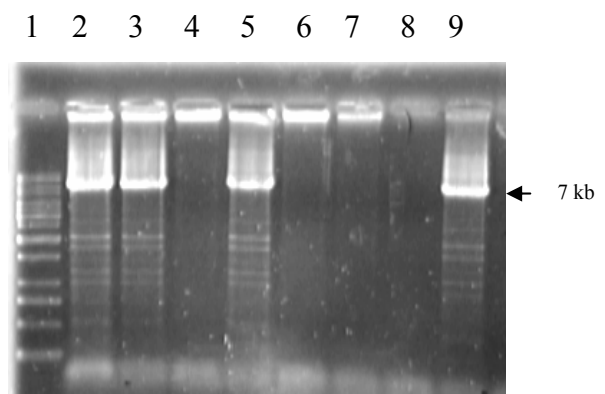
Figure 6.1.14: Genotyping of chimeric mice by PCR. A: a 2.3 kb fragment could be amplified with P366/LacZ-3. from genomic DNA of all these chimeric mice. the marker size was shown on the left side of A; B: similarly, a 7 kb fragment could be amplified with primers P373/P374.

6.1.10 Genotyping of +/- heterozygotes (F1) by PCR and Southern blot

The heterozygotes were genotyped by PCR using the primers P374/P374 as previously. From a total of 146 offspring, 34 (19 male and 15 female) were proven to be heterozygotes (-/+). The fig 6.1.15 shows the PCR products from some offspring.

Figure 6.1.15: Genotyping of heterozygotes by PCR

Lane 1: 1 kb marker
Lane 2, 3, 5: 7 kb fragments from heterozygotes.
Lane 4, 6, 7: no products from these offspring (indicating that they are wild type)
Lane 8: PCR using genomic DNA from wild type mice as a negative control
Lane 9: PCR using genomic DNA from



The genotype of chimeric mice and heterozygotes were further confirmed by Southern blot analysis using an internal probe (part of the mouse ERp28 cDNA including half of the second exon and most of the third exon amplified by primers P294/P278). As we see in

fig 6.1.19, since there is a SpeI site in the sequence of loxneoloxp, an extra SpeI site will be introduced into mouse genomic DNA if homologous recombination has occurred. The hybridisation of the labelled probe prepared above with mouse genomic DNA digested by SpeI should show a single 8.79 kb band in wild type, double bands (8.79/5.4 kb) in +/- mice. As expected, the results shown in fig 6.1.16 further confirmed the PCR results and prove the genotype of the chimera and heterozygotes.

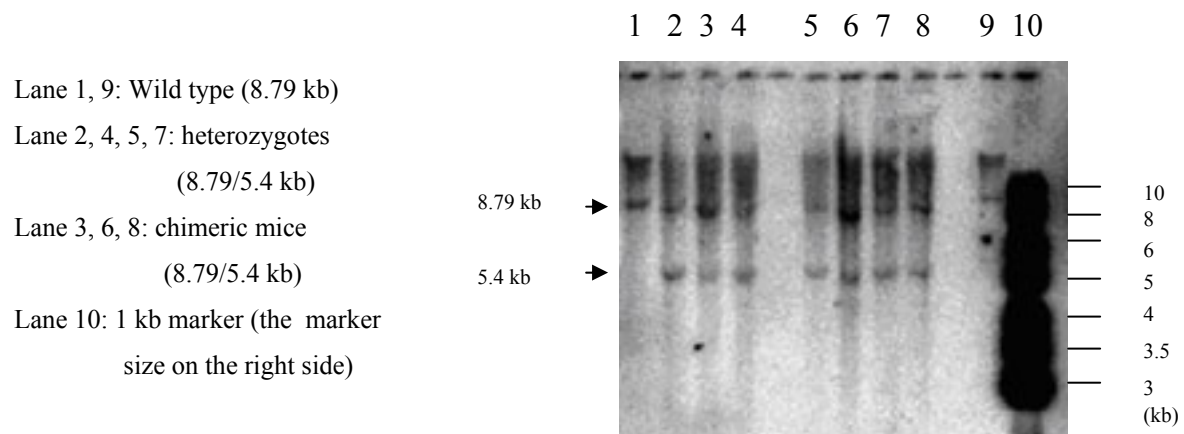


Figure 6.1.16: Genotyping chimeric mice and heterozygotes by Southern blot.

6.1.11. Genotyping of +/- homozygotes (F2) by PCR and Southern blot

Homozygotes were genotyped by PCR and Southern blot. As shown in fig 6.1.17 and fig 6.1.18, PCR was performed with primers P379 (designed from the sequence before the initiator ATG of exon1) and P293 (from exon2) and as expected, a single product with a size of 5 kb was obtained from wild type mice. In contrast, a 7 kb fragment and two products of 5 kb and 7 kb were generated from homozygotes and heterozygotes, respectively. The homozygotes were further confirmed by Southern blot using internal (SpeI/probe II) (fig 6.1.20) and external (HindIII/probe I) probes (fig 6.1.21). The restriction sites (SpeI and HindIII) in the ERp28 locus of wild type mice and in the locus after recombination and the expected Southern blot results are shown in fig. 6.1.19.

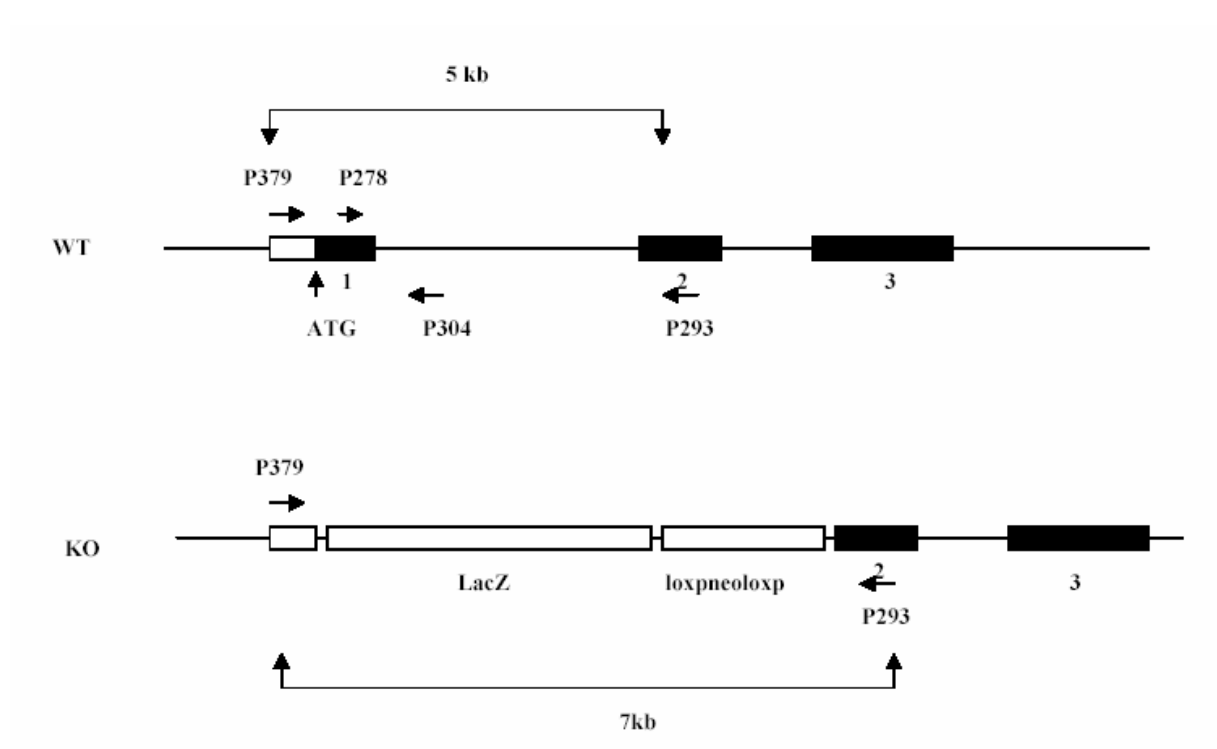


Figure 6.1.17: Primers for genotyping of homozygotes and expected products. Exons are shown as filled boxes.

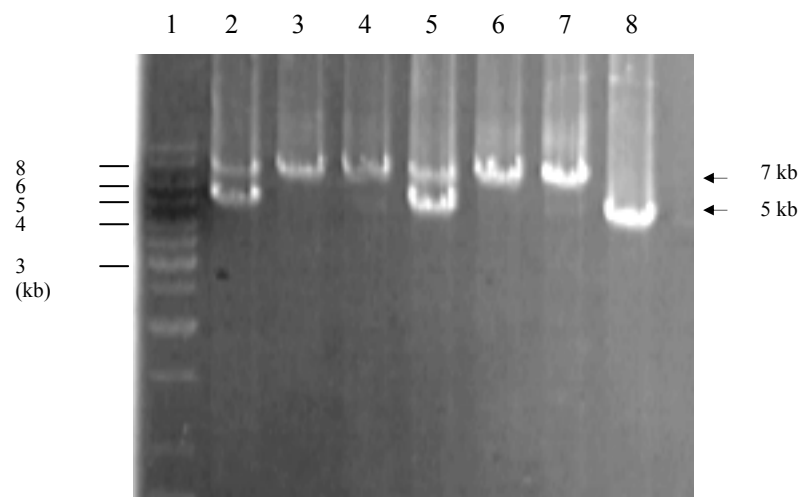
Figure 6.1.18: Genotyping of homozygotes by PCR

Lane 1: marker (the size of the marker is shown on the left side)

Lane 2, 5: heterozygotes (5/7 kb)

Lane 3, 4, 6, 7: homozygotes (7 kb)

Lane 8: wild type (5 kb)



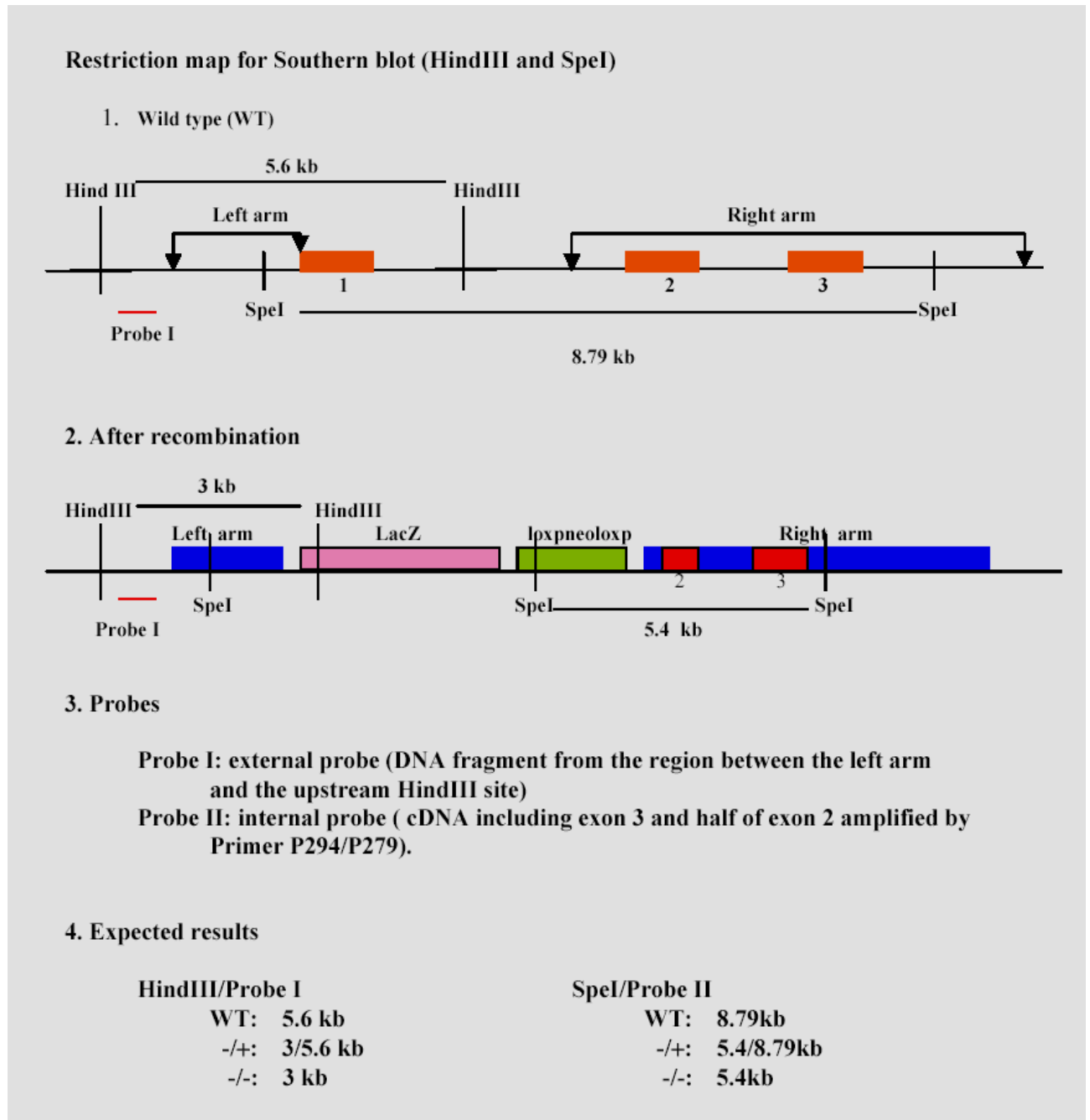


Figure 6.1.19: HindIII and SpeI sites in wild type ERp28 locus and the additional HindIII and SpeI sites after homologous recombination. Probes and expected results. Exons (1-3) are shown as red filled boxes. HindIII and SpeI sites are presented as vertical lines.

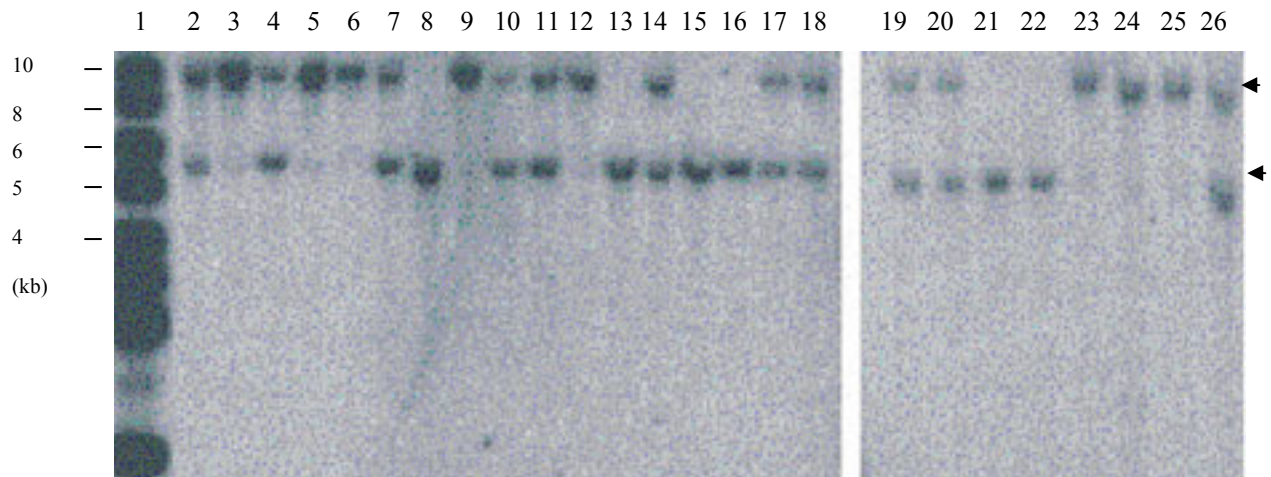


Figure 6.1.20: Genotyping of F2 offspring by Southern blot using an internal probe (SpeI/ probe II).

Lane1: 1 kb marker; Lane 2: chimera 16 as a control (double bands at 8.79/5.4 kb); Lanes 3, 5, 6, 9, 12, 23, 24, 25: wild type mice(+/+) (a single band at 8.79 kb); Lanes 4, 7, 10, 11, 14, 17, 18, 19, 20, 26: heterozygotes(-/+) (double bands at 8.79/5.4 kb); Lanes 8, 13, 15, 16, 21, 22: homozygotes (-/-) (a single band at 5.4 kb). The arrowheads on the right side represent 8.79 kb and 5.4 kb, respectively. The size of 1 kb marker is marked on the left side.

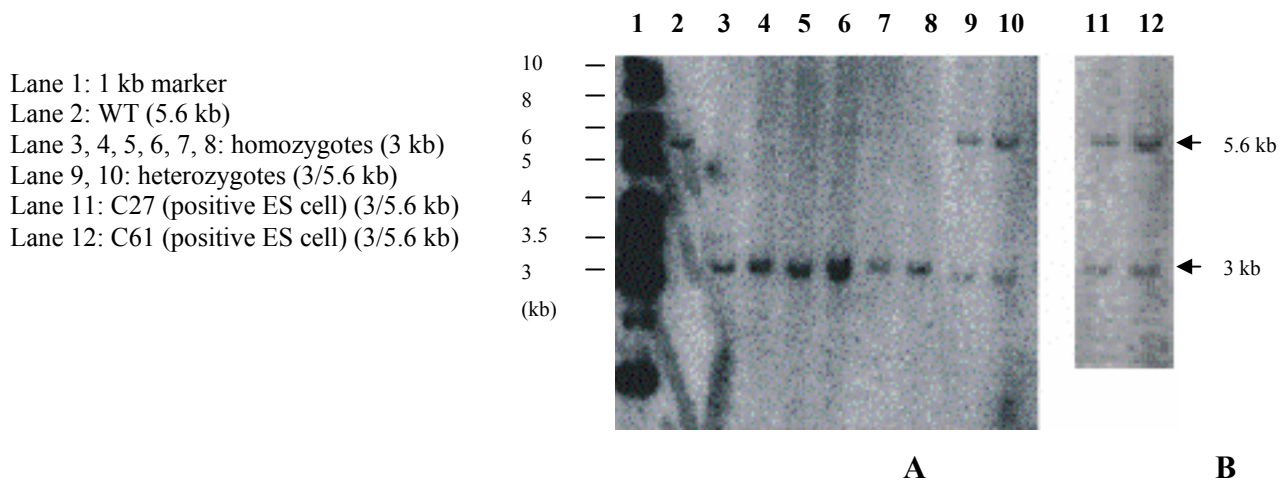
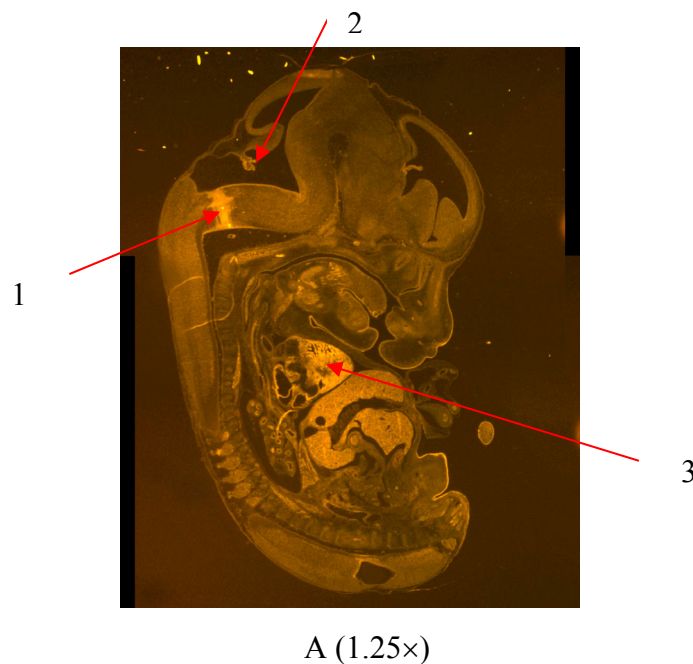


Figure 6.1.21: (A) Genotyping of F2 offspring by Southern blot using an external probe (HindIII/Probe I; (B) The two positive recombinant ES cells (C27 and C61) were confirmed by Probe I. The size of 1 kb marker is marked on the left side. The 5.6 kb and 3 kb signals are marked as arrowheads on the right side.

6.1.12. Immunostaining on paraffin sections of E12 embryos

To check the expression pattern of mouse ERp28, immunostaining was performed on serial paraffin sections of E12 embryos. The sections were stained with anti-ERp28 antibody followed by Cy3-coupled secondary antibody. Control sections were stained only with the secondary antibody under the same conditions. The pictures shown below were taken from sagittal sections. Fig 6.1.22 shows distribution of the staining signals over the whole section. Despite widely distributed weak signals, ERp28 is expressed at relatively high levels in specific regions especially in the radial glial cells of brain stem raphe (medulla oblongata), and the epithelium of the choroid plexus of the fourth ventricle. Due to auto-fluorescence of blood cells, fluorescent signals could be observed even in the heart and liver of the control sections. However, comparing the intensity of fluorescent signals on the sample sections with that on the control section, showed the specific expression of ERp28 in the ventricle of the heart.



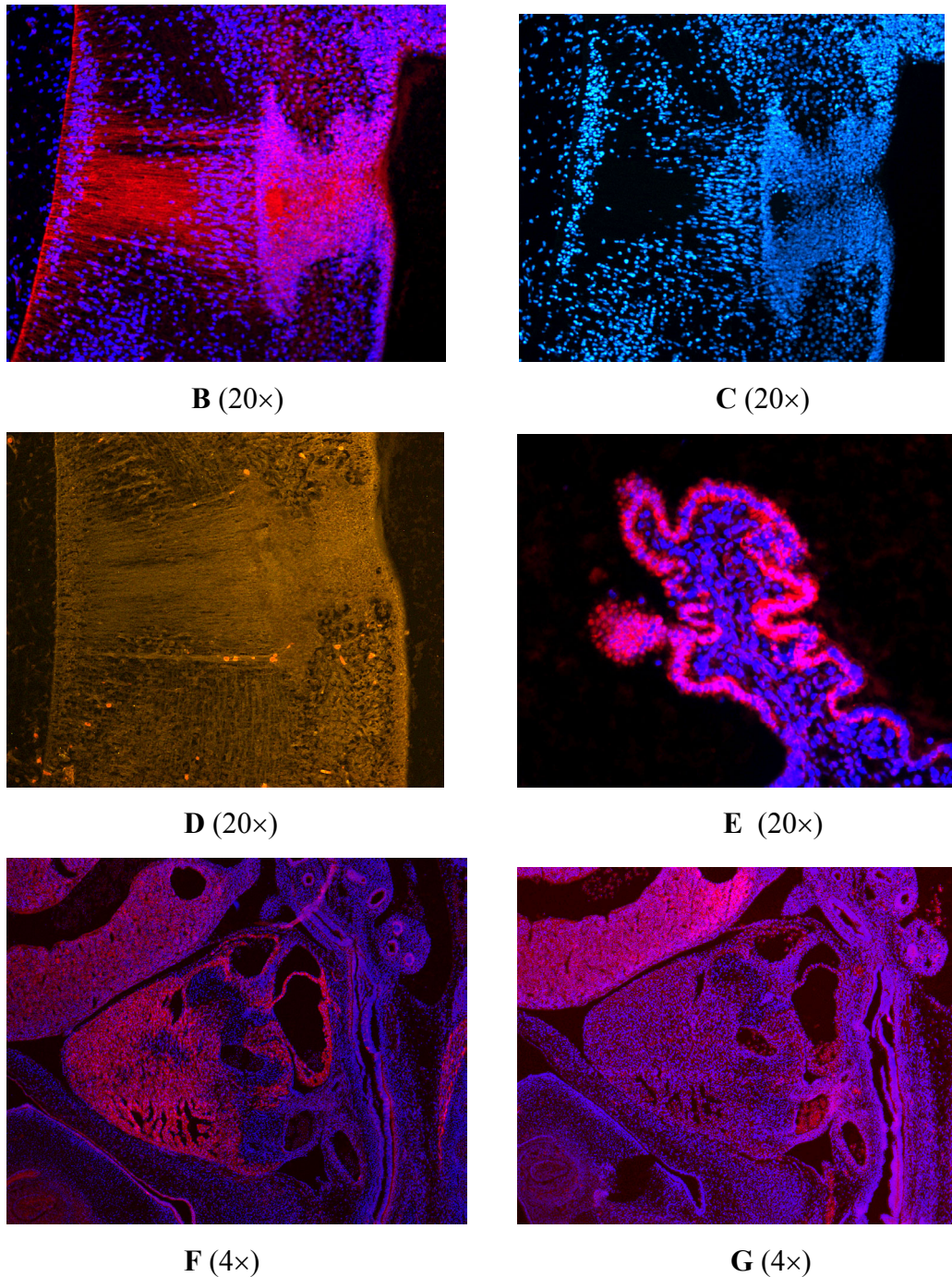


Figure 6.1.22: The expression of ERp28 at stage E12 shown by immunohistological staining. The degree of magnification is given below each figure.

A: whole section (sagittal section through the embryo midline)

B: Signals in radial glial cells of brain stem raphe (red arrow 1 in A), blue is DAPI staining and red is the Cy3 staining of ERp28

C: DAPI staining to show cell nuclei

D: Control section of C (stained only with secondary antibody)

E: Signals (red) in the epithelia of choroid plexuses in the fourth ventricle (red arrow 2 in A), blue is DAPI staining

F: Signals (red) in the heart (red arrow 3 in A) and DAPI staining (blue)

G: Control section of F (stained only with secondary antibody)

6.2 Structural and functional analysis of Wind

6.2.1 SDS-PAGE gel and native gel electrophoresis of purified recombinant proteins His-Wind and Wind-His

Both His-Wind and Wind-His could be expressed in XL1 Blue cell and purified with a high yield (50-100 mg from 1 L culture). The purity of both proteins is quite high (more than 95%) as shown by SDS-PAGE analysis (fig 6.2.1), and sufficient for crystallization and biochemical experiments. On native gel, His-Wind shows a single band with about 55 kD in size, indicating that His-Wind forms a homodimer in solution. Similar results were obtained with Wind-His (data not shown).

Figure 6.2.1: SDS-PAGE gel and Western blot of His-Wind

A: His-Wind on 12.5% SDS-PAGE gel

B: Western blot result using anti-His antibody

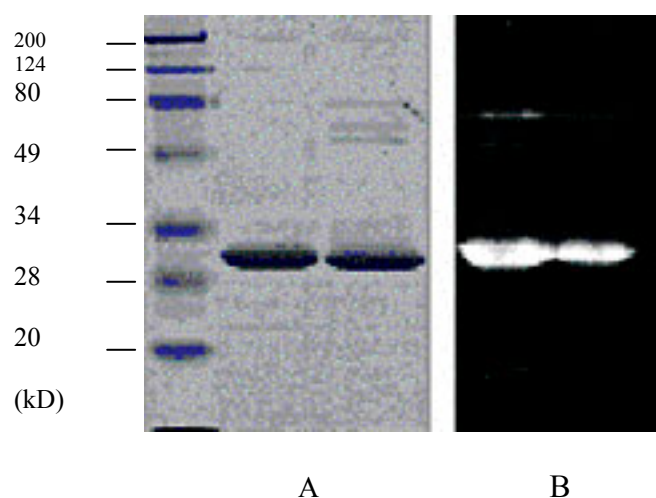
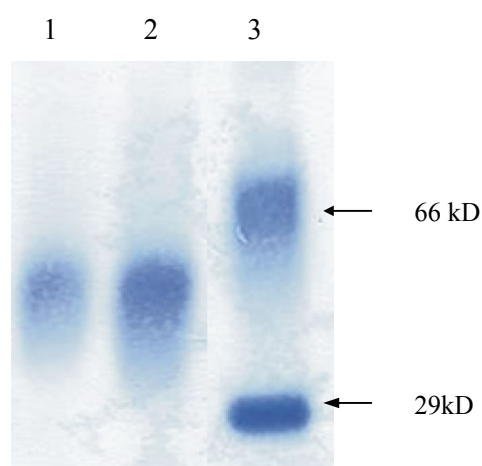


Figure 6.2.2: His-Wind on 7.5% native gel

Lane 1: 5 μ g

Lane 2: 10 μ g

Lane 3: marker



6.2.2 Cross-linking of His-Wind and Wind-His with glutaraldehyde

Cross-linking experiments were performed to confirm the dimerization of Wind. The same amount of His-Wind and Wind-His (about 71 μ g) was cross-linked with glutaraldehyde

(final concentration 1%) in the presence of different concentration of NaCl (0, 50, 150 and 500 mM) at pH 7.4. The results (fig 6.2.3) showed that both proteins can form a dimer of 55 kD. However, under these conditions, Wind-His is more likely to form a dimer than His-Wind, probably due to some inhibition of dimer formation by the N-terminal His-tag. (These experiments were performed by D.M.Ferrari).

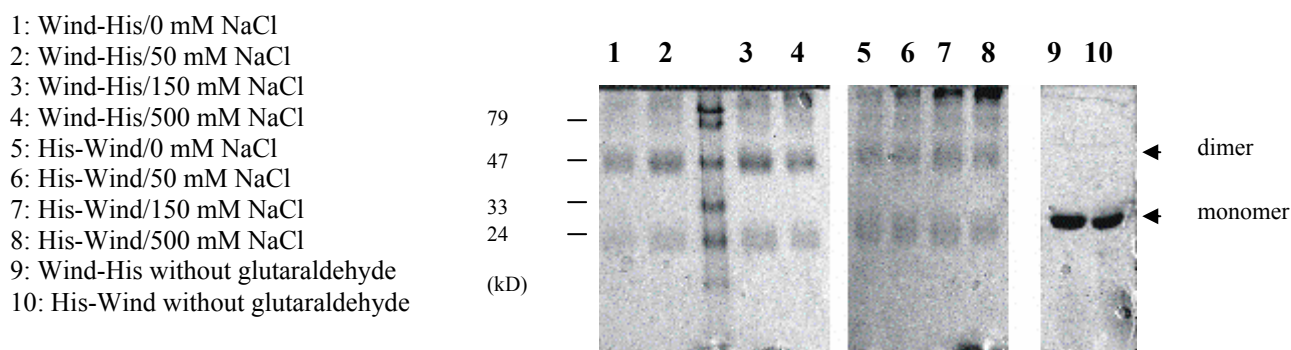


Figure 6.2.3: Cross-linking of His-Wind and Wind-His with glutaraldehyde. The samples were run on 10% SDS-PAGE gel. The size (kD) of the protein marker is shown on the left side. Dimer and monomer of Wind are shown with the arrowheads on the right side.

6.2.3 Insulin reduction assay

The standard insulin reduction assay was performed to test the redox activity of the -CTGC- motif of Wind. As shown in fig 6.2.4, in contrast to the positive control with CaBP2, neither His-Wind nor Wind-His is able to reduce the disulfide bonds of insulin. Clearly, no difference could be observed between buffer, His-Wind and Wind-His.

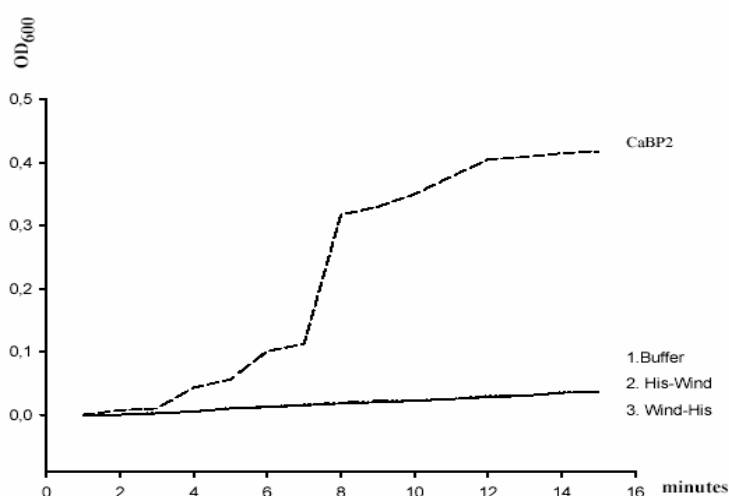


Figure 6.2.4: Insulin reduction assay. The upper line refers to CaBP2. The lower lines represent buffer, His-Wind and Wind-His.

6.2.4 Crystal structure of His-Wind

(The structure analysis was primarily performed and most of the pictures of crystal structure were made in collaboration with Dr. Qingjun Ma, Prof. Dr. Isabel Uson, and Prof. Geroge.M. Sheldrick at the Department of Structural Biology, University of Goettingen at the Department of Structural Chemistry of University of Goettingen)

6.2.4.1 Overall structure

Wind crystallizes as a homodimer (here the two monomers are referred to as A and B respectively). Each of the monomers contains two distinct domains connected by a flexible short linker of 11 residues (residues 140-150). The secondary structure elements are arranged as $\beta 1$ - $\alpha 1$ - $\beta 2$ - $\alpha 2$ - $\beta 3$ - $\alpha 3$ - $\beta 4$ - $\beta 5$ - $\alpha 4$ - $\alpha 5$ - $\alpha 6$ - $\alpha 7$ - $\alpha 8$ - $\alpha 9$. The N-terminal domain from $\beta 1$ to $\alpha 4$ is a thioredoxin-like domain (b domain) with about 117 amino acids (residues 23-139). This domain forms a very characteristic thioredoxin fold which was first observed in the structure of thioredoxin. In this fold, five β strands form the core β sheet with $\beta 4$ antiparallel to the other strands, surrounded by four α helices. The unique C-terminal D-domain is composed of 107 residues and has a 5-helix fold in which all helices are in anti-parallel arrangement. The missing C-terminal tail (including the KEEL signal) is probably disordered as a flexible loop.

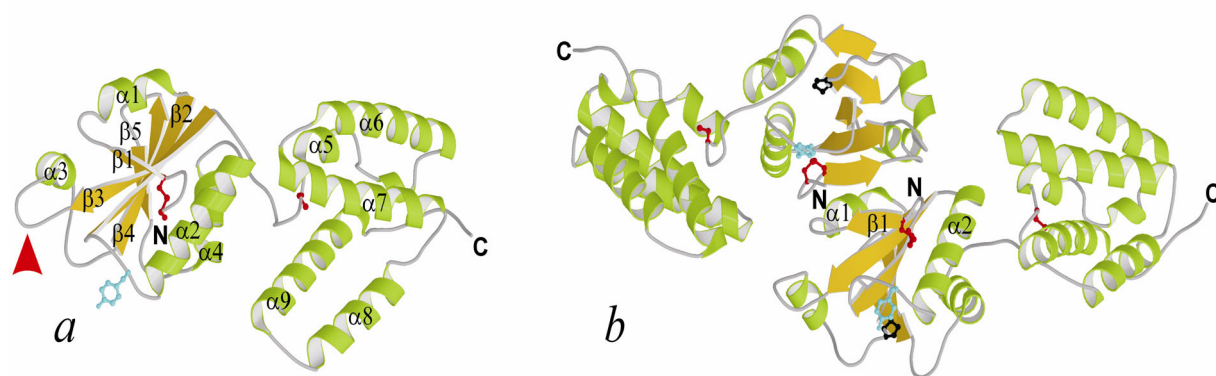


Figure 6.2.5: Structure of the Wind homodimer. (a) Ribbon structure of a Wind monomer. The left represents the thioredoxin domain and the right part is the helical D-domain. The arrow indicates the pentapeptide loop unique to PDI-D β proteins. β strands are in yellow, helices in green. The three cysteines are shown in red, Tyr55 in blue. (b) Ribbon structure of the Wind dimer. The molecule is rotated 90° compared to the monomer in (a). Colors as in (a). Structural elements involved in dimer interaction are labelled. Pro106 is presented as a black stick model. (Image reproduced with permission from Ma *et al.*, 2003. J. Biol. Chem. in press).

6.2.4.2 Dimer structure

The crystal structure of Wind is a homodimer with dimensions of $106 \times 53 \times 37 \text{ \AA}^3$, formed by a head-to-tail arrangement of the N-terminal b-domains. The D-domains are separated by the b-domains and lie on opposite sides. The dimer interface is formed mainly by the residues on either side of $\beta 1$ (residues 24-34), residues within and after $\alpha 1$ (residues 37-43) and residues within and between $\alpha 2$ and $\beta 3$ (residues 67-77) via hydrophobic interactions as well as hydrogen bonding. The buried area in the interface for each monomer is about 758 \AA^2 , corresponding to about 6.3% of the surface of the monomer. The two $\beta 1$ strands are roughly antiparallel, however they interact with each other indirectly through a hydrogen bond network mediated by 5 water molecules rather than through direct interactions to form a continuous beta sheet.

As shown in fig 6.2.8 (A), dimerization generates a deep hydrophilic cleft with approximate dimensions of $11 \times 11 \times 27 \text{ \AA}^3$, between the b domains of the monomers. The residues from the loop between $\beta 2$ and $\alpha 2$ and residues from the end of $\beta 13$ and the following loop flank the cleft (hereafter referred to as the “dimer cleft”). The residues within and around $\beta 1$ are at the cleft’s base.

6.2.4.3 The conserved residues on the protein surface show a distinct pattern

The sequences of four members of the PDI-D β family (Wind, rat ERp29/28, human ERp28/29 and mouse 28/29) were aligned. In the b domain of Wind, the identical amino acids form three major clusters ($V_{47}K_{48}F_{49}D_{50}$ in $\beta 2$, $Y_{53}P_{54}Y_{55}G_{56}E_{57}K_{58}$ in the loop $\beta 2$ - $\alpha 2$, and $\alpha 2$, and DYG in the loop $\beta 3$ - $\alpha 3$). Interestingly (see from fig 6.2.6/6.2.7, it is obvious that on the surface of the protein, most of the identical residues are distributed on the side with the dimer cleft, in contrast to very few on the opposite surface. The two identical clusters (YPYGEK) and ($D_{87}Y_{86}G_{88}$) are exposed on the surface of the b domain. Meanwhile, D50 in the third identical cluster (VKFD), together with P106 preceding $\beta 4$ and $Y_{53}P_{54}Y_{55}G_{56}E_{57}K_{58}$, form a cluster which mimics the cluster around the active site of thioredoxin. On the surface of the D-domain, the identical residues belong mainly to $\alpha 8$ and $\alpha 9$.

Results

		<u>β1</u>	<u>α1</u>	<u>β2</u>	
<i>Drosophila</i>	-----MMHILVTLTLLVAIHSIPT-TWAVTCTGCVDLDELSEKFTVERFPYSVV	47			
rat	--MAAAVPGAVSLSPLLSVLLGLLLLLSAPHGASGLHTKGALPLDVTTFYKVIKSKFVLV	58			
mouse	MAAAAGVSGAASLSPLLSVLLGLLLLLFAPHGGSGLHTKGALPLDVTTFYKVIKSKFVLV	60			
human	--MAAAVPRAAFLSPLLPLLLGFLLLSAPHGGSGLHTKGALPLDVTTFYKVIKSKFVLV	58			
	: : * * . : * . : . * : * : * : * : : * :				
		<u>α2</u>	<u>β3</u>	<u>α3</u>	
<i>Drosophila</i>	KFDIAYPYGEKHEAF TAFSKSAHKATKDLLIATVGVKDYGELENKALGDRYKVDKFNPS	107			
rat	KFDTQYPYGEKQDEFKRLAENS-ASSDDLVAE VGISDYGDKLNME LSEK YKLDKESYPV	117			
mouse	KFDTQYPYGEKQDEFKRLAENS-ASSEELLVAE VGISDYGDKLNME LSEK YKLDKESYPV	119			
human	KFDTQYPYGEKQDEFKRLAENS-ASSDDLVAE VGISDYGDKLNME LSEK YKLDKESYPV	117			
	*** *****: * . : : : : : : : : * * : * : * : * : * : * : * :				
		<u>β4</u>	<u>β5</u>	<u>α4</u>	<u>α5</u>
<i>Drosophila</i>	IFLFGKNADEYVQLPSHVDVTLDNLKAFVSANTPLVIGRDGCIKEFNEVLKNYANIPDAE	167			
rat	FYLFRR-DGDFENPVPYSGAVKVGAIQRWLKG-QGVYLGMPGCLPAYDALAGQFIEASSRE	175			
mouse	FYLFRR-DGDLNPVLYNGAVKVGAIQRWLKG-QGVYLGMPGCLPAYDALAGEFIKASSIE	177			
human	FYLFRR-DGDFENPVPYTGAVKVGAIQRWLKG-QGVYLGMPGCLPVYDALAGEFIRASGVE	175			
	: : * : . * : * . : : : : : : * : * : * : : : : . . . *				
		<u>α6</u>	<u>α7</u>	<u>α8</u>	
<i>Drosophila</i>	QLKLIKQLQAKQEQLTDPEQQQNARAYLIYMRKIHEVGYDFLEETKRLRLKAG-KVTE	226			
rat	ARQAILKQGQDGLSGVKETDKKWA SQYLKIMGKILDQGEDFPASELARISKLIEN-KMSE	234			
mouse	ARQAILKQGQDGLLSVKETEKKWA SQYLKIMGKILDQGEDFPASEMARIGKLIEN-KMSD	236			
human	ARQALLKQGQDNLSVKETQKKWAEQYLKIMGKILDQGEDFPASEMTRIARLIEKNKMSD	235			
	: : * . . : : * * * * : * * * . * * : * * : * : :				
		<u>α9</u>			
<i>Drosophila</i>	AKKEELLRKLNILEVFRVHKVTKTAPEKEEL	257			
rat	GKKEELQRS LNILTAFR-----KKGAEKEEL	260			
mouse	SKKEELQKSLNILTAFR-----KKEAEKEEL	262			
human	GKKEELQKSLNILTAFAQ-----KKGAEKEEL	261			
	. ***** : . ***** . * : * . *****				

Figure 6.2.6: Sequence alignment of PDI-Dβ proteins using ClustalW. Identical residues are red and marked with asterisks. Similar residues are marked with dots. Secondary structures are shown above the sequence and the range is marked with a blue bar. Signal sequences are in blue.

(A): Conserved surface on the top face around the dimer cleft. The molecule is rotated 90° compared to fig 6.2.5 (b). The residues identical in PDI-Dβ proteins (fig 6.2.6) are colored in magenta. The position of the CTGC motif is marked with a yellow circle; P106 with a small green triangle; Y55 with a black arrow.

(B): Conserved surface on the bottom face of Wind. Colour as in (A). The dimer is rotated 180° compared to (A).

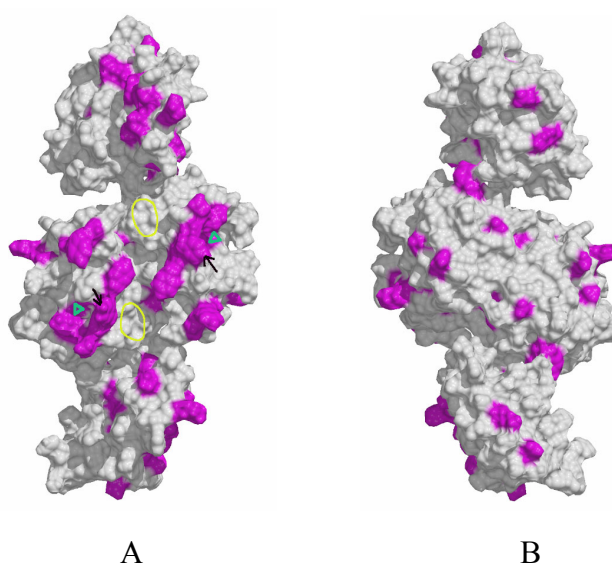


Figure 6.2.7: Conserved surface of Wind.

6.2.4.4 The electrostatic potential and the hydrophobic patches on the surface

A very obvious and interesting feature is the negatively charged dimer cleft. This negative potential is contributed by residues D29, D31, E32 around $\beta 1$ and E88, E90 in the loop $_{\beta 3-\alpha 3}$ from each monomer. However, only D31 is completely conserved in other PDI-D β proteins. The surface of b-domains are slightly negative, while the D-domains show a slightly positive potential. The negative potential in the dimer cleft again suggests that the dimer and dimer cleft may be important for the function of Wind.

Although no large hydrophobic patches are found on the dimer surface, some small patches on top of or beside the dimer cleft form a significant hydrophobic area. Three conserved tyrosines Y53, Y55 and Y86, together with residues I51, A52 and F105 from each monomer significantly contribute to this area.

Figure 6.2.8: Surface electrostatic potential.

(A): Orientation as in fig 6.2.7 (A). Negative potential is colored in red, positive potential in blue.

(B): Surface electrostatic potential. Orientation as in fig 6.2.7 (B). Colored as in (A).

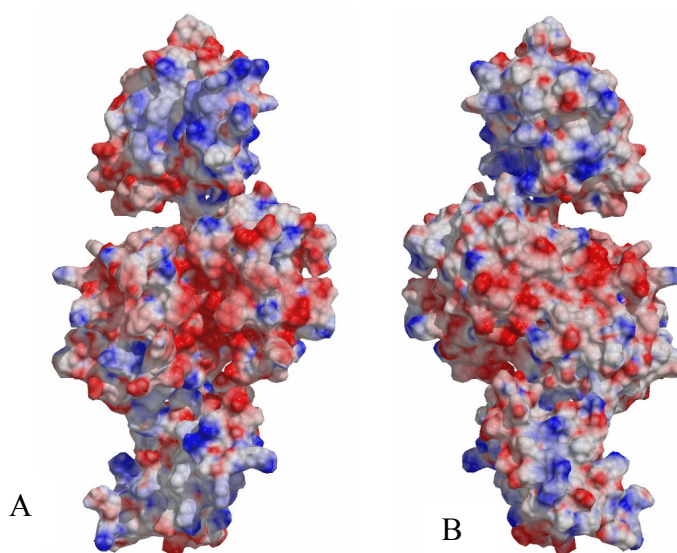
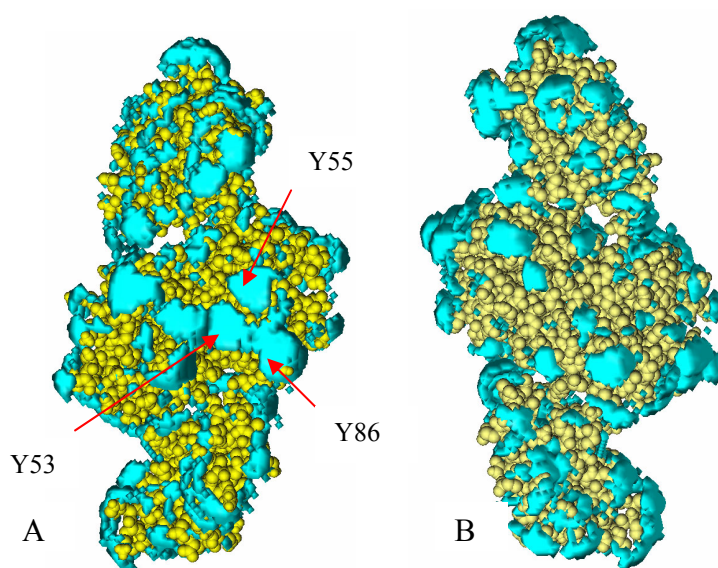


Figure 6.2.9: Surface hydrophobic potential

(A): Hydrophobic patches (blue) are shown as volumes of the hydrophobic potentials contoured at about -2.48 Kcal/mol. Orientation as in fig 6.2.7 (A). The three conserved tyrosines, Y53, Y55, Y86 are marked with red arrows.

(B): Surface hydrophobic potential. Colors as in (A); Orientation as in fig 6.2.7 (B)



6.2.4.5 Other features of the structure

The N-terminal –CTGC- motif forms a disulfide bond in both monomers and both sulfur atoms in this motif are buried inside the protein and are not solvent-accessible. The other conserved cysteine (C149) is located in the linker region and has a free thiol which is solvent-accessible.

Another feature that should be mentioned is the presence of a conserved proline (P106) adopting a less common *cis* conformation, located in the loop preceding β_4 , which is conventionally called the *cis*-Pro loop found in many other TRX-like proteins. It is known that the corresponding *cis*-proline in other TRX-like proteins is important not only for maintaining the local structure in this region but also for their activities (Qin *et al.*, 1996).

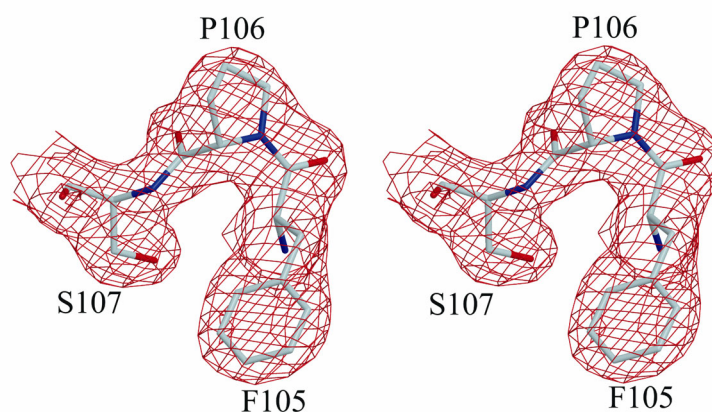


Figure 6.2.10: Stereo view of the *cis*-Pro106. The electron density of F105-S107 are shown.

6.2.4.6 Comparison of Wind and other PDI-related proteins

An alignment based on structure of the b-domains of Wind and ERp29, and the a- and b-domains of PDI is shown in fig 6.2.11. Compared with the PDI a- and b-domains, there are two major differences in Wind: first, the redox active site –CGHC- in the PDI a-domain is replaced by a cluster of identical residues in PDI-D β proteins (YPYGEK); second, a surface exposed pentapeptide insertion (DYGEL) which is partially conserved in PDI- D β proteins is located in the loop $_{\beta_3-\alpha_3}$, a region that is known to allow introduction of insertions in many thioredoxin-fold containing proteins such as DsbA (Martin *et al.*, 1993), glutathione peroxidase (Epp *et al.*, 1983. Martin, 1995), and DsbC (McCarthy *et al.*, 2000).

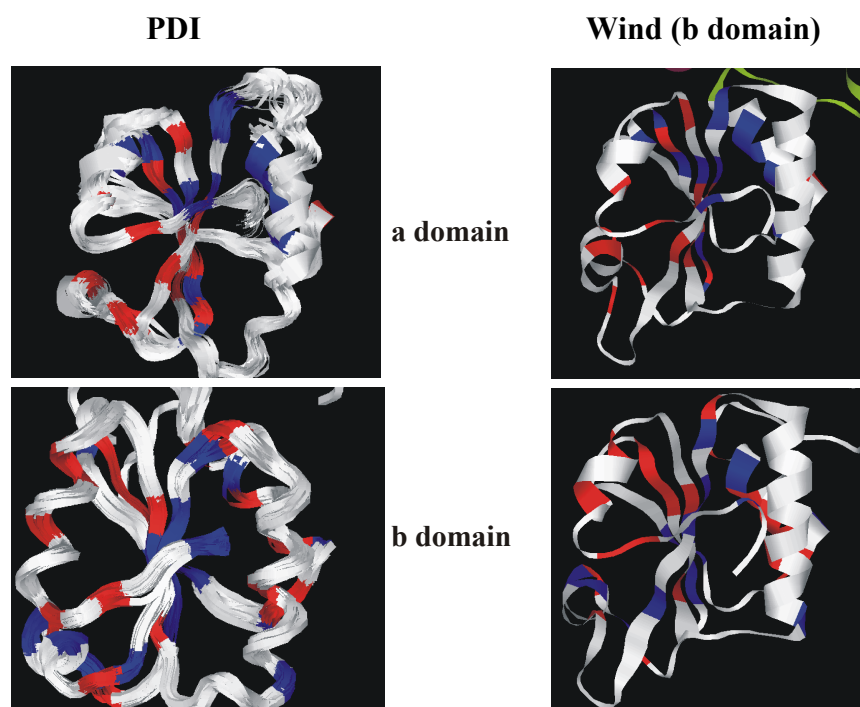


Figure 6.2.12: Identical and similar residues between the b domain of Wind and the a, b domains of PDI in the thioredoxin fold. Identical residues are colored in red, similar residues in blue. (data from D.M.Ferrari).

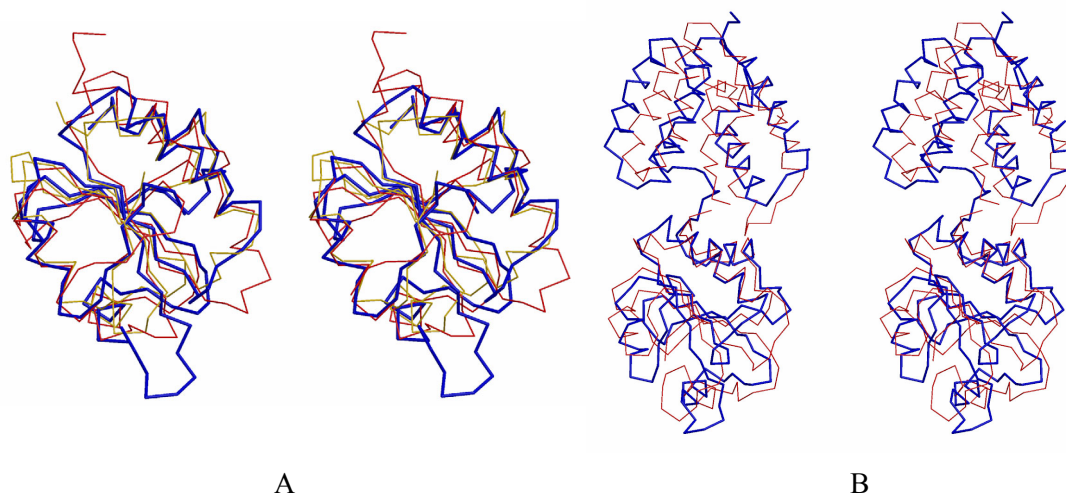


Figure 6.2.13: Stereo view of the superpositions of Wind and other PDI-related proteins. A: Superposition of Wind b-domain and PDI a- & b domain, respectively. Wind colored in blue, PDI a domain in red, PDI b domain in gold. The protruding loop in Wind (at the bottom of the map) is the additional pentapeptide in PDI-D β proteins. B: Superposition of Wind and ERp29. The C α trace of Wind is colored in blue, ERp29 colored in red. The b-domain and the D-domain of ERp29 are superposed onto the corresponding part of Wind, respectively.

7. Discussion

7.1 Studies on ERp28

7.1.1 The gene of ERp28 contains three exons

Since the sequence of the mouse genome of some strains is already known at present, finding out a gene's structure has become an easy task. However, when I started the project of ERp28 knock out mice, the intron/exon structure of ERp28 was still a mystery. The combinations of results from RT-PCR, PCR and the comparison between the cDNA sequence and the known sequences of some genomic clones, revealed that ERp28 contains three exons. The first exon contains the translation start codon ATG after a very short 5' untranslated region (UTR), while the third exon which is also the largest has the translation stop codon TGA before the 3' UTR. The first intron between exon1 and exon 2 is about 5 kb in size, the second intron between exon 2 and exon 3 is about 2.5 kb long.

7.1.2 ERp28 may have an alternative splice form among mammalian cells

The RT-PCR results indicate that at least in mouse brain tissue of E12.5 embryos and in F9 cells, there is a possible alternative splice form of ERp28 in which the second exon is spliced out. Furthermore, searching the EST sequence database showed that the mammalian homologues of mouse ERp28 (human ERp28, rat ERp29) also have this alternative form. Therefore, it is likely that this alternative form is conserved among mammals. In *Drosophila*, *windbeutel*, the homologue of ERp28, contains only a very short intron (about 400 bp) (Konsolaki *et al.*, 1998) and the corresponding region of the second exon of ERp28 does not possess the putative splicing signal GT-AG. This suggests that the alternative splicing form may exist only in a limited number of species.

It is known that alternative splicing can generate multiple transcripts encoding proteins (they may or may not be isoforms) with subtle or opposing functional differences that can have profound biological consequences (Kriventseva *et al.*, 2003; Kondrashov *et al.*, 2003). However, the alternative splice form of ERp28 cannot be translated into a protein if the same translation start point is used because of a reading frame shift caused by the lack of the second exon. This raises several possibilities and questions: (1) the splice form might be translated into a new protein using a different translation start point. For this, only the ATGs in the third exon can be used as the translation start point because the ATGs in the first exon can only give some very short peptides. If this is the case, the alternative splicing form will form a

protein whose amino acid sequence is identical to the C-terminal part of ERp28. The third exon encodes the last part of the thioredoxin domain and the complete D-domain of ERp28 protein. This means that the alternative splice product does not contain a signal sequence and is unlikely to locate to the ER. In addition, EST database (mouse and human) searches indicate that many tissues and cells have this alternative splice form. This indicates that this protein, like ERp28, is widely expressed. However, since the result from Western blotting were not conclusive (data not shown), further experiments need to be performed to verify the existence of this new protein. (2) the other possibility is that the translation start point of the alternative splice form is different from that of ERp28 mRNA, being located further upstream of the one used by ERp28. However, such an ATG could not be found in either rat, mouse or human genomic DNA sequences between the promoter and exon 1 of ERp28. (3) the alternatively spliced mRNA of ERp28 may not be translated into a protein, but might function at the mRNA level. (4) it can not be excluded that this alternative splice form is just a non-functional byproduct. As reported, more than a dozen cancers and inherited diseases in humans (and mice) are associated with abnormalities in alternative splicing (Dredge *et al.*, 2001). However, the alternative splicing of ERp28 is not likely to fit this scenario because it can take place in normal animals and tissues (according to the source of ESTs).

7.1.3 ERp28 is expressed at high level in some tissues at stage E12 of mouse embryo development

The immunostaining results reveal that although ERp28 is widely expressed in mouse embryos at stage E12, in some tissues, it has a relatively higher expression level. However, it should be remembered that these data are preliminary and only from E12 stage sections. Further studies should be carried out to clarify the tissues and cell types in which ERp28 is expressed and to uncover the expression pattern of ERp28 throughout the developmental process of mouse embryos using sections of other stages or by whole mount staining.

The strongest signals are present in the raphe of the brain stem (medulla oblongata). Observation at high magnification clearly shows that the cells in this region have a very distinct morphology which is similar to that of radial glial cells. Radial glial cells are identified by their unique morphology and location in the developing brain. They consist of a bipolar soma, with one short process based on the ventricular surface, and a long radial fibre spanning the cerebral wall to the pial surface. Actually, the recognition of radial glial cell depends more on the morphology than use of marker proteins due to the lack of very specific markers for radial glial cells (Zhang, 2001). Traditionally, radial glial cells are believed to be astrocyte precursors and serve as a scaffold to support and direct neurons during their

migration. However, recent studies have provided strong support for their role as precursor cells in the ventricular zone that generate cortical neurons and glia, in addition to providing migration guidance (Garcia-Verdugo *et al.*, 1998; Chanas-Sacre, 2000; Parnavelas and Nadarajah, 2001; Gregg *et al.*, 2002; Liour and Yu, 2003; Fishell and Kriegstein, 2003). From the morphology, these cells with high expression of ERp28 look like radial glial cells, however it should be further confirmed by staining with relatively specific radial glial cell markers such as RC2 (radial cell 2) (Misson *et al.*, 1988 a, b; Zhang, 2001).

The second tissue with stronger signals are the choroid plexuses of the fourth and lateral ventricles. At higher magnification, we can clearly see that the signals are present in the epithelial cells of choroid plexus. Mammalian choroid plexuses develop at four sites in the roof of the neural tube shortly after its closure, in the order IVth, lateral, and IIIrd ventricles (K.M. Dziegielewsaka, 2001). Choroid plexuses have two major functions: (1) they form one of the blood-brain barrier interfaces that control the brain's internal environment. The mechanisms involved combine a structural diffusion restraint (tight junctions between the plexus epithelial cells) and specific exchange mechanisms. (2) the choroidal epithelium is the major producer of cerebrospinal fluid (CSF). The secretion of CSF is fulfilled by a process that involves the movement of Na^+ , Cl^- and HCO_3^- from the blood to the ventricles of the brain. This creates an osmotic gradient, which drives the secretion of H_2O . Meanwhile, some proteins may also be transferred from the blood to the CSF by the choroid plexus epithelium. In addition, the choroid plexus epithelium secretes many proteins and polypeptides (Speake *et al.*, 2001; Chodobski and Szmydynger-Chodobska, 2001). Compared with that of progenitor cells, the ependymal cells, the signal of ERp28 in the epithelium of the choroid plexus is much higher. This is consistent with the fact that ERp28 usually has a higher expression level in secretory cells or tissues (Demmer *et al.*, 1997; Shnyder and Hubbard, 2002). In addition, it should be noted that at the stage E12, the choroid plexuses of the fourth and lateral ventricles are just becoming evident. This suggests that ERp28 is likely to be expressed in the epithelial cells of the choroid plexuses even at the beginning of their differentiation from ependymal cells (Sarnat, 1998).

Due to the auto-fluorescence of blood cells, fluorescence signals could also be observed even in heart and liver of control sections. However, these signals are distributed evenly. In the sample sections, the signal intensity in the heart is much stronger than that of the control section and the ratio of the signal intensity between heart and liver or other tissues is much higher than that of the control section. This indicates that ERp28 is expressed in the heart (especially in the ventricles) at stage of E12.

In addition, some other tissues or organs, for example, tongue muscle, lung and kidney also express ERp28 at stage E12. Since a LacZ reporter gene was introduced into the genome of knock out mice, it will be much easier to analyse the expression pattern using the x-gal staining method in future experiments.

7.1.4 The phenotype of ERp28 knock-out mice, a mystery to be resolved

The aim of a gene knock out in mice is to elucidate the function by analyzing the phenotype of the knock out mice. Therefore, the phenotype analysis is the final and crucial step. Although females homozygous for *wind* mutations can survive, grow normally and produce eggs that have normal morphology, the deletion of *windbeutel* in *Drosophila* causes the loss of dorsal-ventral patterning of the embryos and eventually leads to lethality (Konsolaki and Schupbach, 1998). A putative reason for this phenotype is that Wind is essential for the translocation of Pipe to the Golgi through a chaperone/escort function. Without Wind, Pipe cannot exit the ER and thus cannot fulfill its function as a proteoglycan modifying enzyme (Sen *et al.*, 2000). A possible phenotype of ERp28 deletion mice might be similar to Wind deficient *Drosophila*, if ERp28 is also involved in the folding and/or transport of the mammalian homologues of Pipe, heparan sulfate 2-O sulfotransferase. The deletion of ERp28 might then cause the loss of the function of heparan sulfate 2-O sulfotransferase. Heparan sulfate (HS) is a long unbranched polysaccharide found covalently attached to various proteins at the cell surface and in the extracellular matrix. It plays a central role in embryonic development and cellular function by modulating the activities of an extensive range of growth factors and morphogens (Selleck, 2000). HS 2-O-sulfotransferase (Hs2st) occupies a critical position in the succession of enzymes responsible for the biosynthesis of HS, catalysing the transfer of sulfate to the C2-position of selected hexuronic acid residues within the nascent HS chain. Previous studies have concluded that 2-O-sulfation of HS is essential for its involvement in many growth factor/receptor interactions. Surprisingly, Hs2st^{-/-} mice made by the gene trap method did not show general tissue or organ defects as expected. However, although offspring lacking functional Hs2st can survive until birth, they die perinatally due to a complete failure to form kidneys. In addition, Hs2st^{-/-} mice also exhibit defects in the eye and the skeleton (Bullock *et al.*, 1998; Merry and Wilson, 2002). If the relationship between ERp28 and Hs2st is the same as or similar to that between Wind and Pipe, the deletion of ERp28 might lead to similar phenotypes.

However, the homozygotes that we obtained did not exhibit any obvious phenotype so far, at least no similar phenotype to Hs2st knock-out mice, indicating that the deletion of ERp28

in mice does not have severe effects on the function of Hs2st. This suggests that Hs2st is either not be a substrate of ERp28 or the function of ERp28 can be compensated by other genes. On the other hand, since the severe effect of the deletion of *windbeutel* is shown on *Drosophila* embryos, without obvious effect on the development of female homozygotes and male homozygotes, the consequence of the functional loss of ERp28 may first become obvious in the next generation of knock out mice.

The rat homologue, ERp29 has been reported to be a component of thyroglobulin /chaperone complexes (Sargsyan *et al.*, 2002). Expression of ERp29 bearing Q70A or C157A point mutations in the rat thyroid cell line resulted in the drastic reduction of thyroglobulin secretion (Mkrtchian, *et al.*, 2003; personal communication). Since mouse ERp28 is highly similar to rat ERp29, we could speculate that in the ERp28^{-/-} mice, the folding or secretion of thyroglobulin should be affected as well.

Because of the high expression level in the radial glial cells and the epithelium of choroid plexus, these cells will also be a focus for further analysis. However, as the phenotype of gene knock-out mice is not always consistent with the expression pattern of the gene in question, the defects of knock out mice may take place in other less obvious tissues or organs.

It should be noted that if the protein product of the alternative splice form exists, the generated knock-out mice will disrupt both the functions of ERp28 and those of the alternative splice form and the phenotype will represent the loss of functions of both. To distinguish the functions of them, an alternative splice form-specific knock out mouse would be helpful (Hammes *et al.*, 2001).

7.2 Studies on Wind

7.2.1 Wind has two distinct domains and forms a homodimer

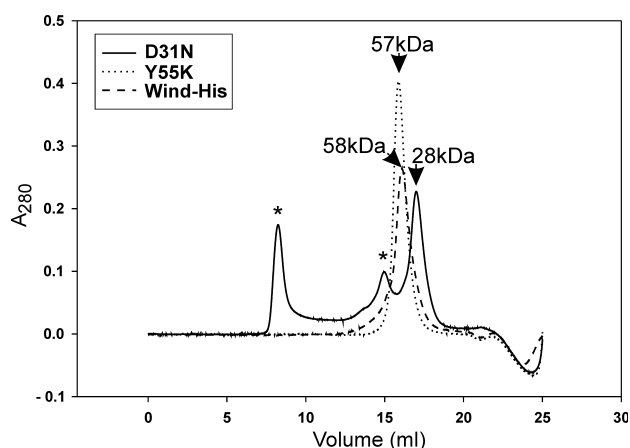
The expressed His-Wind in *E. coli* has a molecular weight around 27 kD as expected. The crystal structure of Wind clearly shows a dimer and each monomer contains two distinct domains, the N-terminal thioredoxin-like b domain and the C-terminal α helical D-domain connected by a flexible linker of 11 amino acids. The dimer interface was contributed only by the b domain without the participation of the D domain. It is very probable that in solution Wind also exists as a dimer. This point is supported by several assays: Wind runs as a single band with the size of 2×27 kD on the native gel; This 2×27 kD is further confirmed by the size exclusion chromatography, performed by D.M Ferrari using a Superdex 200 HR column (fig.7.2.1). Cross-linking experiments (performed by D.M Ferrari) of the purified His-Wind and Wind-His show that Wind can form a dimer, although significant amounts of monomer

are also present (in the presence of 50-150 mM NaCl, the ratio (dimer:monomer is roughly 2:1). However, some of the monomer may be due to the low efficiency of the cross-linker under the conditions used. From the above results, we can say that Wind is dimerized both in the crystal and in solution.

Figure 7.2.1: Oligomeric state of Wind

(reproduced with permission from Ma *et al*, 2003. J. Biol. Chem. in press)

Wild type His-tagged Wind (Mw: 58 kDa) has a molecular mass very similar to Y55K mutant Wind (57 kDa), in contrast to D31N mutant Wind (28 kDa). This indicates that Wind forms a dimer in solution and Y55K mutant does not change the dimerization of Wind, whereas, D31N mutation significantly disrupts the dimerization of Wind. Aggregation could be observed for the D31N mutant (asterisks)



monomer A

Residue	Interface ASA (\AA^2)	H-Bonds
CYS 24	1.93	.
THR 25	30.67	.
GLY 26	35.86	1
CYS 27	8.27	.
VAL 28	63.74	.
ASP 29	22.29	.
LEU 30	2.52	.
ASP 31	37.81	1
LEU 33	100.10	.
SER 34	37.08	.
LYS 37	32.32	.
THR 38	29.91	.
ARG 41	130.93	2
PHE 42	63.04	.
PRO 43	21.79	.
LYS 67	10.55	.
HIS 70	64.11	.
THR 73	9.42	.
LYS 74	35.42	1
ASP 75	4.24	.
LEU 76	6.11	.

monomer B

Residue	Interface ASA(\AA^2)	H-Bonds
CYS 24	1.58	.
THR 25	24.16	.
GLY 26	39.54	1
CYS 27	5.92	.
VAL 28	65.08	.
ASP 29	23.10	.
LEU 30	3.45	.
ASP 31	37.88	1
LEU 33	90.29	.
SER 34	35.60	.
LYS 37	97.69	.
THR 38	30.61	.
ARG 41	129.62	1
PHE 42	56.91	.
PRO 43	16.17	.
LYS 67	12.38	.
HIS 70	51.24	.
LYS 71	3.92	.
LYS 74	26.92	1
ASP 75	11.70	1
LEU 77	1.80	.

Table 7.2.1: Residues in the dimer interface as suggested by evaluation of x-ray crystallography. (a):Residues in monomer A. (b):Residues in monomer B (the monomers A and B is referred to different mainchain in the dimer, respectively). ASA: accessible surface area. The most important residues for the interface are shadowed either in yellow (those forming hydrogen bonds) or pink (those having highly buried area).

The dimer interface of Wind is different from that of the recently reported ERp29 dimer revealed by NMR studies (Liepinsh *et al*, 2001). In Wind, the most important residues involved in the interface are G26, V28, D31, L33, H37, R41, F42, H70, H74, and D75. These residues can either form hydrogen bonds or are highly buried. They are located in β 1, within or after α 1, within or after α 2. In contrast, the dimer interface suggested for rat ERp29 is on the opposite side of the TRX-like domain and mainly composed by residues around β 5, loop $_{\beta$ 5- α 5, and α 5. To confirm that the dimer interface of Wind in solution is the same as in the crystal, a D31N mutant was made by D.M Ferrari in our group that should probably disrupt the dimerization based on the crystal structure. *In vitro* experiments (multiangle light scattering of protein separated by size exclusion chromatography and cross-linking) show that D31N has a much smaller tendency to dimerize (fig.7.2.1) (Ma *et al.*, 2003. J. Biol. Chem in press). This indicates that the dimer interface derived from the crystal structure exists in solution as well. Generally, homologues should have the same or similar structures. The differences between the dimer interface of Wind and rat ERp29 were unexpected. From our data, it clearly shown that the dimer of Wind exists both in the crystal and in solution and the dimer interface should be correct. Two possibilities can explain this discrepancy, either the dimer interface of rat ERp29 as given by Liepinsh *et al* is not correct or a different interface really exists in rat ERp29. In fact, in recent experiment, it could be shown that mutating D42 (the residue corresponding to D31 in Wind) in human ERp28 to N significantly inhibit dimer formation, indicating that ERp28 may indeed have a dimer interface similar (if not identical) to Wind. (experiments by D.M. Ferrari, data not shown).

7.2.2 Is the dimerization functional?

The cross-linking results of overexpressed Wind in COS7 cells (data not shown) have indicated that, as its mammalian homologues ERp28/29, Wind can form dimers as well as monomers and higher oligomers. As known, Wind is a dimer both in the crystal and in solution. A question we have asked is whether the dimer is essential for the function of Wind. From the crystal structure, we can see that the dimerization of Wind generates a significant cleft that is large enough to hold a small peptide. Interestingly, the cleft is hydrophilic and negatively charged. This fact lets us to speculate that the negative charge in the cleft may be necessary for the substrate binding. If so, it would be analogous to the mechanism suggested for the interaction of the sarcoplasmic reticulum protein calsequestrin with the membrane protein Junctin (Zhang *et al.*, 1997). The putative substrate of Wind, Pipe, indeed has a region (S96-N139) with a high content of basic residues (pI=11.3) which is distinct from the more C-

terminally located sequences with net negative charge. Actually, one of the possible active-sites of Pipe (identified by sequence alignment with other sulfotransferases) is in the middle of this positive charged region. If this speculation is true, then we can imagine that the dimerization of Wind would be necessary for its function, as otherwise the negatively charged cleft would not be formed. In addition, since the residues that contribute to the negative potential of the dimer cleft are not conserved in ERp28/29 and there are no basic residue rich regions in mammalian heparan sulfate 2-O sulfotransferase and other isoforms of Pipe, this negative dimer cleft may provide the functional specificity of Wind.

However, the mutant D31N that could disrupt the dimerization significantly *in vitro* did not show obvious effects on the transport of Pipe to the Golgi *in vivo* (fig.7.2.3). This may be due to residual dimer that could not be disrupted completely *in vivo*. As indicated by cross-linking experiments with intact cells, it seems less likely that dimerization is not necessary for the function. This very interesting and important question needs to be finally clarified in the future.

7.2.3 Both the b and D domain are necessary for the localization of Pipe

Wind contains two distinct domains, b and D. The b domain is necessary for dimerization while the D-domain is not as shown in the crystal structure. We asked ourselves what the respective functions of the b- and D-domains could be and whether both domains are necessary for the localization of Pipe. To answer these questions, two truncated Wind constructs (Wind-N and its identical construct fused to a KEEL retention/retrieval sequence, Wind-N-KEEL) which only contain the N-terminal b domain, were co-expressed with Pipe in COS7 cells. Neither Wind-N nor Wind-N-KEEL was able to translocate Pipe to the Golgi although Wind-N-KEEL showed a clear ER distribution (fig.7.2.3, experiments by K. Barnewitz). This indicates that the D-domain also participates either in the folding or the transport of Pipe. The necessity of the D-domain was also supported in a Wind mutant fly. Two mutations resulting in severe developmental defects were identified in the D domain (Konsolaki and Schupbach, 1998). The *wind*^{M46} allele contained a single base change generating a stop code after Q-187 and leading to a truncated D-domain in which nearly all of the last three α -helices were missing. The *wind*^{T6} allele had a single base change in the NIL motif, resulting in the conversion of L-239 to a proline. The L-to-P change might disrupt the structure of helix α_9 although L239 is located at the very end of α_9 .

Interestingly, replacing the D-domain of Wind with that of mouse ERp28 (Wind-N-p28D) did not affect the capability of Wind to process Pipe although the full length mouse ERp28 failed to promote the translocation of Pipe to the Golgi (fig.7.2.3).

These results indicate that both the b- and D-domains are indispensable for the translocation of Pipe. However, as the function of the D-domain of Wind can be replaced by that of mouse ERp28, we can imagine that the D domains in both proteins play a very similar role and the functional specificity of Wind must be determined by the b domain.

7.2.4 The –CTGC- motif is redox inactive and not involved in the localization of Pipe

In contrast to ERp28/29, Wind possesses a –CTGC- motif at the very end of the N-terminus. From the crystal structure, we can clearly see that this motif is not located in the position of the classical redox-active –CXXC- site of the thioredoxin fold of thioredoxin and other members of PDI-family (normally it exists between the C-terminus of the loop β_2 - α_2 and the beginning of α_2). However, the position of the –CXXC- motif and the thioredoxin fold are not always essential for the redox activity as reported that some proteins having –CXXC- motifs but without a TRX-like fold (Langenbach and Sottile, 1999; O’Neil *et al.*, 2000). Pipe, the putative substrate of Wind, contains 7 cysteines in total. Four of them are conserved in other 2-O sulfotransferases and they may form disulfide bonds. These facts prompt us to ask whether the –CTGC- motif is redox active and necessary for the localization of Pipe.

Using the standard insulin reduction assay, His-Wind was not able to reduce the disulfide bonds in insulin. Since the –CTGC- motif is at the very beginning of the Wind sequence and since, from the structure, we could see that the –CTGC- motif is at the bottom of the dimer cleft, the additional N-terminal His-tag sequences (MRGSHHHHHHGS) of the recombinant protein may have affected the activity of the –CTGC- motif. However, the C-terminal His-tagged Wind (Wind-His), in which the –CTGC- motif should be well exposed, showed the same result, indicating that this motif is not redox active. Moreover, Wind could not complement the activity of PDI in yeast, indicating that Wind lacks a general catalytic redox/isomerase activity (Ma *et al.*, 2003. J.Biol.Chem in press).

Furthermore, a Wind mutant in which the –CTGC- motif was changed to STGS did not show any effects on the localization of Pipe compared with wild type Wind (fig.7.2.3). This suggests that at least in this process, the function of Wind does not require the –CTGC- motif and redox/isomerase activity. Therefore we can speculate that Wind may function as a redox-independent chaperone/escort protein.

In *Anopheles*, a homologue of Wind and ERp28 was discovered recently (EAA06716). The first cysteine of the –CTGC- motif of Wind is lost, in contrast, the second cysteine is still maintained. Considering that the closer relationship between *Anopheles* and *Drosophila* than that between *Drosophila* and mammals, we can not exclude the possibility that the –CTGC- motif is just an evolutionary residue without important function. Another interesting point is that although the identity of the b domain of the *Anopheles* homologue (compared to that of Wind) is higher than for mammalian ERp28/29 (43% in *Anopheles*-ERp28, 35% in mammalian ERp28), the D-domain of Wind shows the same degree of similarity with the *Anopheles* ERp28 (31%) as with the mammalian ERp28 (30%).

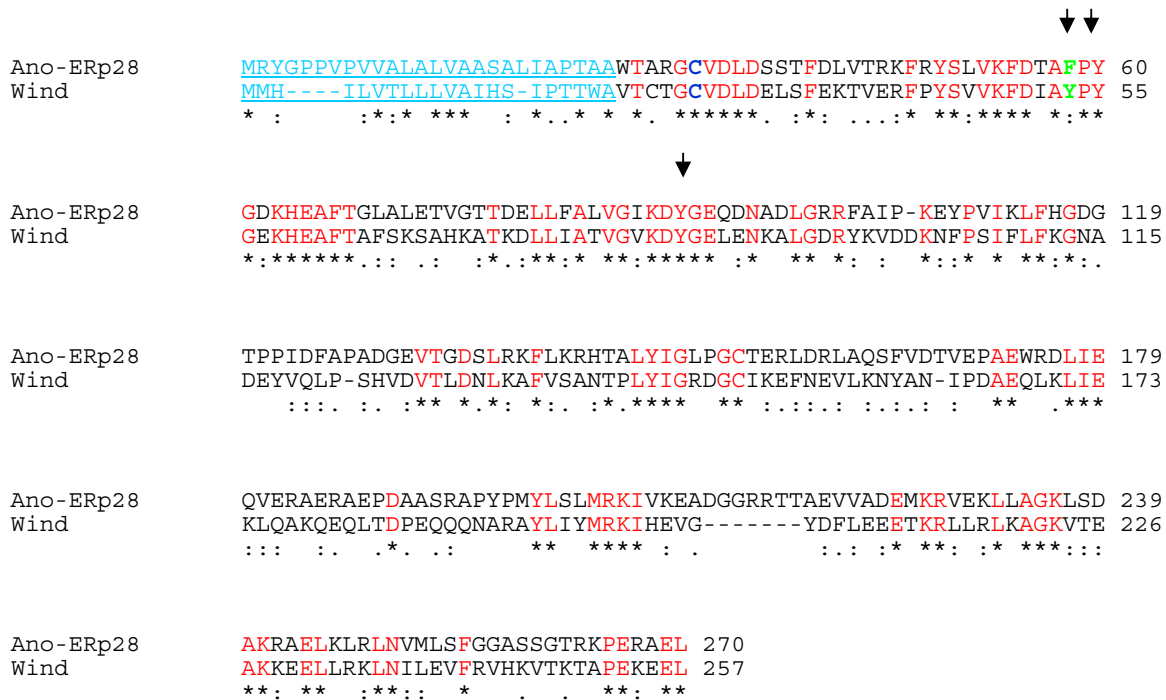


Figure 7.2.2: Alignment of Anopheles-ERp28 (Ano-ERp28) with Wind. The identical residues are shown in red and marked with stars below. Similar residues are marked below with dots. The three tyrosines (Y53, Y55, Y86) are marked by arrows on the top. Y53 in Wind and the corresponding residue F in Anopheles-ERp28 are colored in green.

7.2.5 The possible substrate binding site on the surface of the b domain

The conserved residues of PDI-D β homologues show a very characteristic distribution on the surface of the crystal structure (fig.6.2.7). The surface around the cleft, which includes a loose cluster of three tyrosines (Y53, Y55, and Y86) shows considerable conservation, suggesting that the surface may be involved in substrate binding. Y55, which is surface exposed at the verge of the dimer cleft, was replaced with lysine. The resultant mutant (Wind-Y55K) which is expected to disrupt the surface features, failed to translocate Pipe to

the Golgi (fig.7.2.3), indicating that the conserved surface around Y55 and/or the surface of the dimer cleft formed by the b domain is required for Pipe export from the ER. In fact, our preliminary data suggest that mutations of Y53 and Y86 also disable Wind to translocate Pipe. Furthermore, *in vitro*, we can show that the failure of the Y55K mutant is not due to the disruption of dimerization because multiangle light scattering assays showed that the recombinant protein of Y55K has a molecular mass 57 kD, corresponding to the size of the homodimer of Wind (Ma *et al.*, 2003).

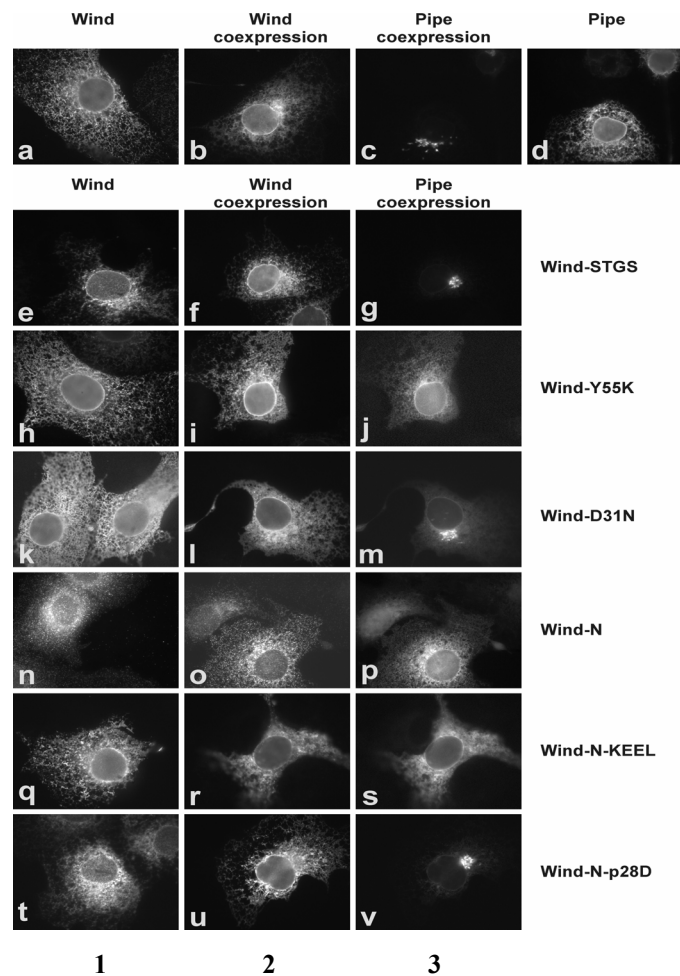


Figure 7.2.3: Effects on the transport of Pipe of different forms of Wind (mutanted or truncated). Pictures in column 1 show the expression of different Wind mutants alone; Column 2: the expression of different Wind mutants in COS7 cells cotransfected with Pipe; Column 3: the expression and localization of Pipe in the cotransfected cells. (details in text). (image reproduced with permission from Ma *et al.*, 2003. J. Biol. Chem in press).

The next question is whether the interaction is mainly a hydrophobic interaction or whether the polar group (-OH) is required as well, as in the *Anopheles* homologue, where Y53 in Wind is replaced by F. It is likely that the two homologues perform similar functions, because of the similarity between them. If the replacement of Y53 with F does not affect the

function of Wind, it will indicate that these tyrosines (at least Y53) may bind to the substrate without involvement of the polar group (-OH). Another interesting point is whether the *Anopheles* homologue could fulfill the function of Wind in folding and/or transport of Pipe. If so, we could limit the indispensable residues for the function of Wind to regions conserved between Wind and Anopheles-ERp28.

Another interesting point is the position of the three tyrosines which are probably involved in substrate binding. Y53 (F58 in Ano-ERp28) and Y55 are located in the conserved cluster YPYGEK which corresponds to the -CXXC- motif in thioredoxin or the a domain of PDI. Y86 is in the conserved pentapeptide insertion located in the loop β_3 - α_3 . In the thioredoxin fold, this region is a hotspot for insertion of amino acids or domains, for example in DsbA (Martin *et al.*, 1993; Martin, 1995), DsbC (McCarthy *et al.*, 2000) and the newly discovered ER protein, ERp18 (Alanen *et al.*, 2003).

7.2.6 The structure and the possible function of the D domain

As the D domain of Wind could be replaced with that of ERp28 without interfering with the function of Wind, it is likely that the D domains in both proteins adopt similar conformation or structure and have similar function. Furthermore, this raises the question whether the D domains in other homologues can also replace the function of the D-domain of Wind. If so, these D domains may possess general function and these functions may be performed by similar mechanisms.

However, the conformation of the D-domain of rat ERp29 is different from that of Wind (fig.6.2.13 B). If the same conformation of ERp29 was adopted by Wind, some hydrophobic residues would be exposed on the surface, which is unlikely to occur, in general. Moreover, the D-domain of Wind displays some flexibility suggested by the difference in B-factors of 26.9Å between the D domains of the two monomeric subunits and by the relative movement of the two domains, amounting to more than 30°, seen when the two domains are superimposed. Taking into consideration the low overall similarity between Wind and ERp29 (44% similarity, 33% identity), some species-specific differences in the structure of the D-domain cannot be ruled out (Ma *et al.*, 2003. J. Biol. Chem in press).

As reported, the truncated Wind without the C-terminal KEEL retention/retrieval sequence (Sen *et al.*, 200) and the *Dictyostelium* homologue, Dd-PDI which have no KEEL at all can be efficiently retained in the ER (Monnat *et al.*, 2000), indicating that the D-domain may carry out a retention/retrieval function and possibly that this function is independent of

the conformation or structure. Indeed, the last 57 C-terminal residues of Dd-PDI can sufficiently localize a green fluorescent protein (GFP) chimera to the ER (Monnat *et al.*, 2000). However, the indispensability of the D-domain for Wind in translocation of Pipe to the Golgi indicates that the retention/retrieval of protein in the ER is not the only function of the D-domain and that it could also take part in substrate binding or processing although it is not clear how this happens.

The flexible linker between the b-domain and the D-domain provides the possibility that the D domain could move relative to the b domain. This movement may be important for substrate binding. When the substrate is present, the two D domains in the homodimer may move to the corresponding b domain, and thus may provide another binding region or mechanism to hold the substrate in addition to the surface of the b-domain. Similar domain movement is also observed in DsbA (Guddat *et al.*, 1998) and DsbC (Mccarthy *et al.*, 2000) structures, which may be important for their functions. In order to show the importance of the corresponding movement between the D-domain and b-domain, it is worthwhile finding out the effects on the function of Wind by shortening the linker region or mutating important residues (Y143, I144, G145) to decrease the flexibility of the linker.

Another point that should be mentioned is the internal KEEL motif which is completely conserved in mammalian homologues (ERp28/29). Interestingly, in *Anopheles* homologue both the C-terminal and internal KEEL are replaced by RAEL. From the similarity, we can assume that the C-terminal RAEL is for ER localization. Whether the internal KEEL motif is important for function/localization of Wind remains to be established.

As discussed above, the D-domain is necessary for the translocation of Pipe and it may provide a binding site for Pipe. If so, then where could the binding site be? Based on the finding that the D domain of Wind could be replaced by that of mouse ERp28, we can imagine that the conserved residues should be the important candidates. Mutation of these residues may provide meaningful information.

7.2.7 Comparison of the b domain of Wind with the a- and b-domains of PDI

Compared with the a domain of PDI, the b-domain of Wind has two major differences. First, the CGHC motif of the a domain is replaced by a potential substrate binding site (Y₅₅ in the conserved residue cluster, YPYGEK) of Wind. Second, in the b domain of Wind, a conserved pentapeptide insertion is located in the loop β_3 - α_3 . On the other hand, they share some common features. Although the sequence similarity is not high, they all adopt the same characteristic thioredoxin fold. As in the a domain but not the b domain of PDI, the b domain

of Wind also contains a *cis*-proline (fig.6.2.10/11) which is assumed to form part of a hydrophobic substrate-binding site close to the active site in thioredoxin-like proteins (Nordstrand *et al.*, 1999), and its *cis* conformation is important for maintaining the integrity of the thioredoxin fold. In addition, the conserved residues between the b-domain of Wind and the a- and b-domains of PDI are primarily within the β -strand elements of the fold or on buried faces of helices (fig.6.2.12), indicating that they are likely to serve a mainly structural role.

7.2.8 Wind provides very useful information for the study of ERp28

The research on Wind can provide much important information for studying and understanding the functions of ERp28. First, the involvement of Wind in the dorsal-ventral patterning of *Drosophila* embryos provides a hint that ERp28 may have the same or similar functions. Second, the result that the D-domain of Wind can be replaced by that of mouse ERp28 indicates that both D domains make the same contribution to the protein functions. Third, since the possible binding surface of Wind is conserved in ERp28/29, we can imagine that ERp28 may have the same substrate binding surface although at present we do not know the mechanism that controls the interaction specificity.

Finally, the structure of Wind reported here is the first crystal structure of a PDI-related protein of the ER to be determined. A closer analysis of the peptide-binding properties of Wind may reveal important clues on peptide-binding sites and mechanisms of function of other PDI proteins.

References:

- Akagi S, Yamamoto A, Yoshimori T, Masaki R, Ogawa R, Tashiro Y. (Localization of protein disulfide isomerase on plasma membranes of rat exocrine pancreatic cells. *J Histochem Cytochem.* 1988 Aug;36(8):1069-74.
- Alanen HI, Williamson RA, Howard MJ, Lappi AK, Jantti HP, Rautio SM, Kellokumpu S, Ruddock LW. Functional characterization of ERp18, a new endoplasmic reticulum-located thioredoxin superfamily member. *J Biol Chem.* 2003 Aug 1;278(31):28912-20.
- Amiri A, Stein D. Dorsoventral patterning: a direct route from ovary to embryo. *Curr Biol.* 2002 Aug 6;12(15):R532-4.
- Anelli T, Alessio M, Mezghrani A, Simmen T, Talamo F, Bachi A, Sitia R. ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family. *EMBO J.* 2002 Feb 15;21(4):835-44.
- Bader M, Muse W, Ballou DP, Gassner C, Bardwell JC. Oxidative protein folding is driven by the electron transport system. *Cell.* 1999 Jul 23;98(2):217-27.
- Bader M, Muse W, Zander T, Bardwell J. Reconstitution of a protein disulfide catalytic system. *J Biol Chem.* 1998 Apr 24;273(17):10302-7.
- Bader MW, Xie T, Yu CA, Bardwell JC. Disulfide bonds are generated by quinone reduction. *J Biol Chem.* 2000 Aug 25;275(34):26082-8.
- Barbouche R, Miquelis R, Jones IM, Fenouillet E. Protein-disulfide isomerase-mediated reduction of two disulfide bonds of HIV envelope glycoprotein 120 occurs post-CXCR4 binding and is required for fusion. *J Biol Chem.* 2003 Jan 31;278(5):3131-6.
- Bardwell JC. Building bridges: disulphide bond formation in the cell. *Mol Microbiol.* 1994 Oct;14(2):199-205.
- Bardwell JC. Disulfide bond formation, a race between FAD and oxygen. *Dev Cell.* 2002 Dec;3(6):758-60
- Bardwell JC, Lee JO, Jander G, Martin N, Belin D, Beckwith J. A pathway for disulfide bond formation in vivo. *Proc Natl Acad Sci U S A.* 1993 Feb 1;90(3):1038-42.
- Bardwell JC, McGovern K, Beckwith J. Identification of a protein required for disulfide bond formation in vivo. *Cell.* 1991 Nov 1;67(3):581-9.
- Benham AM, Cabibbo A, Fassio A, Bulleid N, Sitia R, Braakman I. The CXXCXXC motif determines the folding, structure and stability of human Ero1-Lalpha. *EMBO J.* 2000 Sep 1;19(17):4493-502.
- Bergeron JJ, Brenner MB, Thomas DY, Williams DB. Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *Trends Biochem Sci.* 1994 Mar;19(3):124-8.
- Bose S, Weikl T, Bugl H, Buchner J. Chaperone function of Hsp90-associated proteins. *Science.* 1996 Dec 6;274(5293):1715-7.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976 May 7;72:248-54.

Bu G. The roles of receptor-associated protein (RAP) as a molecular chaperone for members of the LDL receptor family. *Int Rev Cytol.* 2001;209:79-116.

Bu G, Geuze HJ, Strous GJ, Schwartz AL. 39 kDa receptor-associated protein is an ER resident protein and molecular chaperone for LDL receptor-related protein. *EMBO J.* 1995 May 15;14(10):2269-80

Bu G, Schwartz AL. RAP, a novel type of ER chaperone. *Trends Cell Biol.* 1998 Jul;8(7):272-6.

Bullock SL, Fletcher JM, Beddington RS, Wilson VA. Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. *Genes Dev.* 1998 Jun 15;12(12):1894-906.

Cai H, Wang CC, Tsou CL. Chaperone-like activity of protein disulfide isomerase in the refolding of a protein with no disulfide bonds. *J Biol Chem.* 1994 Oct 7;269(40):24550-2.

Chanas-Sacre G, Rogister B, Moonen G, Leprince P. Radial glia phenotype: origin, regulation, and transdifferentiation. *J Neurosci Res.* 2000 Aug 15;61(4):357-63.

Chapman, R. Sidrauski, C. and Walter, P: Intracellular signaling from the endoplasmic reticulum to the nucleus. *Annu. Rev. Cell Dev. Biol.* 14:459–85 (1998).

Charbonnier JB, Belin P, Moutiez M, Stura EA, Quemeneur E. On the role of the cis-proline residue in the active site of DsbA. *Protein Sci.* 1999 Jan;8(1):96-105.

Cheung PY, Churchich JE. Recognition of protein substrates by protein-disulfide isomerase. A sequence of the b' domain responds to substrate binding. *J Biol Chem.* 1999 Nov 12;274(46):32757-61.

Chivers PT, Laboissiere MC, Raines RT. The CXXC motif: imperatives for the formation of native disulfide bonds in the cell. *EMBO J.* 1996 Jun 3;15(11):2659-67.

Chodobski A, Szmydynger-Chodobska J. Choroid plexus: target for polypeptides and site of their synthesis. *Microsc Res Tech.* 2001 Jan 1;52(1):65-82.

Cresswell P, Bangia N, Dick T, Diedrich G. The nature of the MHC class I peptide loading complex. *Immunol Rev.* 1999 Dec;172:21-8.

Cuozzo JW, Kaiser CA. Competition between glutathione and protein thiols for disulphide-bond formation. *Nat Cell Biol.* 1999 Jul;1(3):130-5.

Cunnea PM, Miranda-Vizuet A, Bertoli G, Simmen T, Damdimopoulos AE, Hermann S, Leinonen S, Huikko MP, Gustafsson JA, Sitia R, Spyrou G. ERdj5, an endoplasmic reticulum (ER)-resident protein containing DnaJ and thioredoxin domains, is expressed in secretory cells or following ER stress. *J Biol Chem.* 2003 Jan 10;278(2):1059-66.

Dai S, Schwendtmayer C, Schurmann P, Ramaswamy S, Eklund H. Redox signaling in chloroplasts: cleavage of disulfides by an iron-sulfur cluster. *Science*. 2000 Jan 28;287(5453):655-8.

Darby NJ, Creighton TE. Catalytic mechanism of DsbA and its comparison with that of protein disulfide isomerase. *Biochemistry*. 1995 Mar 21;34(11):3576-87. (a)

Darby NJ, Creighton TE. Characterization of the active site cysteine residues of the thioredoxin-like domains of protein disulfide isomerase. *Biochemistry*. 1995 Dec 26;34(51):16770-80. (b)

Darby NJ, Creighton TE. Functional properties of the individual thioredoxin-like domains of protein disulfide isomerase. *Biochemistry*. 1995 Sep 19;34(37):11725-35.(c)

Darby NJ, Penka E, Vincentelli R. The multi-domain structure of protein disulfide isomerase is essential for high catalytic efficiency. *J Mol Biol*. 1998 Feb 13;276(1):239-47.

Delom F, Mallet B, Carayon P, Lejeune PJ. Role of extracellular molecular chaperones in the folding of oxidized proteins. Refolding of colloidal thyroglobulin by protein disulfide isomerase and immunoglobulin heavy chain-binding protein. *J Biol Chem*. 2001 Jun 15;276(24):21337-42.

Demmer J, Zhou C, Hubbard MJ. Molecular cloning of ERp29, a novel and widely expressed resident of the endoplasmic reticulum. *FEBS Lett*. 1997 Feb 3;402(2-3):145-50.

Desilva MG, Lu J, Donadel G, Modi WS, Xie H, Notkins AL, Lan MS. Characterization and chromosomal localization of a new protein disulfide isomerase, PDIp, highly expressed in human pancreas. *DNA Cell Biol*. 1996 Jan;15(1):9-16.

Desilva MG, Notkins AL, Lan MS. Molecular characterization of a pancreas-specific protein disulfide isomerase, PDIp. *DNA Cell Biol*. 1997 Mar;16(3):269-74.

Dredge BK, Polydorides AD, Darnell RB. The splice of life: alternative splicing and neurological disease. *Nat Rev Neurosci*. 2001 Jan;2(1):43-50.

Dziegielewska KM, Ek J, Habgood MD, Saunders NR. Development of the choroid plexus. *Microsc Res Tech*. 2001 Jan 1;52(1):5-20.

Eckert DM, Kim PS. Mechanisms of viral membrane fusion and its inhibition. *Annu Rev Biochem*. 2001;70:777-810.

Edwards MA, Yamamoto M, Caviness VS Jr. Organization of radial glia and related cells in the developing murine CNS. An analysis based upon a new monoclonal antibody marker. *Neuroscience*. 1990;36(1):121-44.

Eklund H, Ingelman M, Soderberg BO, Uhlin T, Nordlund P, Nikkola M, Sonnerstam U, Joelsson T, Petratos K. Structure of oxidized bacteriophage T4 glutaredoxin (thioredoxin). Refinement of native and mutant proteins. *J Mol Biol*. 1992 Nov 20;228(2):596-618.

Ellgaard, L and Helenius, A. ER quality control: towards an understanding at the molecular level. *Current Opinion in Cell Biology*. 13:431–437 (2001)

Ellgaard, L and Helenius, A. Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol.* 4(3):181-91 (2003).

Ellis RJ, van der Vies SM, Hemmingsen SM. The molecular chaperone concept. *Biochem Soc Symp.* 1989;55:145-53.

Epp O, Ladenstein R, Wendel A. The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. *Eur J Biochem.* 1983 Jun 1;133(1):51-69.

Essex DW, Li M. Protein disulphide isomerase mediates platelet aggregation and secretion. *Br J Haematol.* 1999 Mar;104(3):448-54.

Essex DW, Li M, Miller A, Feinman RD. Protein disulfide isomerase and sulfhydryl-dependent pathways in platelet activation. *Biochemistry.* 2001 May 22;40(20):6070-5.

Farwell AP, Lynch RM, Okulicz WC, Comi AM, Leonard JL. The actin cytoskeleton mediates the hormonally regulated translocation of type II iodothyronine 5'-deiodinase in astrocytes. *J Biol Chem.* 1990 Oct 25;265(30):18546-53.

Fassio A, Sitia R. Formation, isomerisation and reduction of disulphide bonds during protein quality control in the endoplasmic reticulum. *Histochem Cell Biol.* 2002 Feb;117(2):151-7.

Feng W, Bedows E, Norton SE, Ruddon RW. Novel covalent chaperone complexes associated with human chorionic gonadotropin beta subunit folding intermediates. *J Biol Chem.* 1996 Aug 2;271(31):18543-8.

Ferrari DM, Nguyen Van P, Kratzin HD, Soling HD. ERp28, a human endoplasmic-reticulum-luminal protein, is a member of the protein disulfide isomerase family but lacks a CXXC thioredoxin-box motif. *Eur J Biochem.* 1998 Aug 1;255(3):570-9.

Ferrari DM, Soling HD. The protein disulphide-isomerase family: unravelling a string of folds. *Biochem J.* 1;339 (Pt 1):1-10 (1999) .

Fewell SW, Travers KJ, Weissman JS, Brodsky JL. The action of molecular chaperones in the early secretory pathway. *Annu Rev Genet.* 2001;35:149-91.

Fishell G, Kriegstein AR. Neurons from radial glia: the consequences of asymmetric inheritance. *Curr Opin Neurobiol.* 2003 Feb;13(1):34-41.

Frand AR, Cuozzo JW, Kaiser CA. Pathways for protein disulphide bond formation. *Trends Cell Biol.* 2000 May;10(5):203-10.

Frand AR, Kaiser CA. The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. *Mol Cell.* 1998 Jan;1(2):161-70.

Frand, A. R. Kaiser, C. A. Ero1p oxidizes protein disulfide isomerase in a pathway for disulfide bond formation in the endoplasmic reticulum. *Mol Cell.* 1999 Oct;4(4):469-77.

Freedman RB, Hirst TR, Tuite MF. Protein disulphide isomerase: building bridges in protein folding. *Trends Biochem Sci.* 1994 Aug;19(8):331-6.

Freedman, R. B. Klappa, P. Ruddock, L. W. Protein disulfide isomerases exploit synergy between catalytic and specific binding domains. *EMBO Rep.* 3(2):136-40 (2002).

Freskgard PO, Bergenheim N, Jonsson BH, Svensson M, Carlsson U. Isomerase and chaperone activity of prolyl isomerase in the folding of carbonic anhydrase. *Science.* 1992 Oct 16;258(5081):466-8.

Gallina A, Hanley TM, Mandel R, Trahey M, Broder CC, Viglianti GA, Ryser HJ. Inhibitors of protein-disulfide isomerase prevent cleavage of disulfide bonds in receptor-bound glycoprotein 120 and prevent HIV-1 entry. *J Biol Chem.* 2002 Dec 27;277(52):50579-88.

Garcia-Verdugo JM, Doetsch F, Wichterle H, Lim DA, Alvarez-Buylla A. Architecture and cell types of the adult subventricular zone: in search of the stem cells. *J Neurobiol.* 1998 Aug;36(2):234-48.

Gerber J, Muhlenhoff U, Hofhaus G, Lill R, Lisowsky T. Yeast ERV2p is the first microsomal FAD-linked sulfhydryl oxidase of the Erv1p/Alrp protein family. *J Biol Chem.* 2001 Jun 29;276(26):23486-91.

Gilbert HF. Protein disulfide isomerase. *Methods Enzymol.* 1998;290:26-50.

Goldberger, R. F., Epstein, C. J., and Anfinsen, C. B. *J. Biol. Chem.* 238, 628-635 (1963).

Gregg CT, Chojnacki AK, Weiss S. Radial glial cells as neuronal precursors: the next generation? *J Neurosci Res.* 2002 Sep 15;69(6):708-13.

Gunther R, Srinivasan M, Haugejorden S, Green M, Ehbrecht IM, Kuntzel H. Functional replacement of the *Saccharomyces cerevisiae* Trg1/Pdi1 protein by members of the mammalian protein disulfide isomerase family. *J Biol Chem.* 1993 Apr 15;268(11):7728-32.

Hammes A, Guo JK, Lutsch G, Leheste JR, Landrock D, Ziegler U, Gubler MC, Schedl A. Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell.* 2001 Aug 10;106(3):319-29.

Hampton, R.Y. ER-associated degradation in protein quality control and cellular regulation. *Current Opinion in Cell Biology.* 14:476–482 (2002).

Harding, H. P. Calton, M. Urano, F. Novoa, I and Ron, D. Transcriptional and translational control in the mammalian unfolded protein response. *Annu. Rev. Cell Dev. Biol.* 18:575–99 (2002).

Hawkins HC, Blackburn EC, Freedman RB. Comparison of the activities of protein disulphide-isomerase and thioredoxin in catalysing disulphide isomerization in a protein substrate. *Biochem J.* 1991 Apr 15;275 (Pt 2):349-53.

Hayano T, Kikuchi M. Molecular cloning of the cDNA encoding a novel protein disulfide isomerase-related protein (PDIR). *FEBS Lett.* 1995 Sep 25;372(2-3):210-4.

Hebert DN, Foellmer B, Helenius A. Calnexin and calreticulin promote folding, delay oligomerization and suppress degradation of influenza hemagglutinin in microsomes. *EMBO J*. 1996 Jun 17;15(12):2961-8.

Helmann JD. OxyR: a molecular code for redox sensing? *Sci STKE*. 2002 Nov 5;2002(157):PE46.

Hendrick JP, Hartl FU. Molecular chaperone functions of heat-shock proteins. *Annu Rev Biochem*. 1993;62:349-84.

High S, Lecomte FJ, Russell SJ, Abell BM, Oliver JD. Glycoprotein folding in the endoplasmic reticulum: a tale of three chaperones? *FEBS Lett*. 2000 Jun 30;476(1-2):38-41.

Hiniker A, Bardwell JC. Disulfide bond isomerization in prokaryotes. *Biochemistry*. 2003 Feb 11;42(5):1179-85.

Hogg PJ. Disulfide bonds as switches for protein function. *Trends Biochem Sci*. 2003 Apr;28(4):210-4.

Holmgren A. Thioredoxin. *Annu Rev Biochem*. 1985;54:237-71.

Holmgren A. Thioredoxin and glutaredoxin systems. *J Biol Chem*. 1989 Aug 25;264(24):13963-6.

Holmgren A. Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide. *Structure*. 1995 Mar 15;3(3):239-43.

Hong CC, Hashimoto C. An unusual mosaic protein with a protease domain, encoded by the nudel gene, is involved in defining embryonic dorsoventral polarity in *Drosophila*. *Cell*. 1995 Sep 8;82(5):785-94.

Hoshijima K, Metherall JE, Grunwald DJ. A protein disulfide isomerase expressed in the embryonic midline is required for left/right asymmetries. *Genes Dev*. 2002 Oct 1;16(19):2518-29.

Hosoda A, Kimata Y, Tsuru A, Kohno K. JPDI, a novel endoplasmic reticulum-resident protein containing both a BiP-interacting J-domain and thioredoxin-like motifs. *J Biol Chem*. 2003 Jan 24;278(4):2669-76.

Huber-Wunderlich M, Glockshuber R. A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Fold Des*. 1998;3(3):161-71.

Hwang C, Sinskey AJ, Lodish HF. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science*. 1992 Sep 11;257(5076):1496-502.

Jakob U, Muse W, Eser M, Bardwell JC. Chaperone activity with a redox switch. *Cell*. 1999 Feb 5;96(3):341-52.

Jordan A, Reichard P. Ribonucleotide reductases. *Annu Rev Biochem*. 1998;67:71-98.

Kassenbrock CK, Garcia PD, Walter P, Kelly RB. Heavy-chain binding protein recognizes aberrant polypeptides translocated in vitro. *Nature*. 1988 May 5;333(6168):90-3.

Katti SK, LeMaster DM, Eklund H. Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution. *J Mol Biol*. 1990 Mar 5;212(1):167-84.

Katzen F, Beckwith J. Transmembrane electron transfer by the membrane protein DsbD occurs via a disulfide bond cascade. *Cell*. 2000 Nov 22;103(5):769-79.

Kemmink J, Darby NJ, Dijkstra K, Nilges M, Creighton TE. Structure determination of the N-terminal thioredoxin-like domain of protein disulfide isomerase using multidimensional heteronuclear ¹³C/¹⁵N NMR spectroscopy. *Biochemistry*. 1996 Jun 18;35(24):7684-91.

Kemmink J, Darby NJ, Dijkstra K, Scheek RM, Creighton TE. Nuclear magnetic resonance characterization of the N-terminal thioredoxin-like domain of protein disulfide isomerase. *Protein Sci*. 1995 Dec;4(12):2587-93.

Kemmink J, Dijkstra K, Mariani M, Scheek RM, Penka E, Nilges M, Darby NJ. The structure in solution of the b domain of protein disulfide isomerase. *J Biomol NMR*. 1999 Apr;13(4):357-68.

Kikuchi M, Doi E, Tsujimoto I, Horibe T, Tsujimoto Y. Functional analysis of human P5, a protein disulfide isomerase homologue. *J Biochem (Tokyo)*. 2002 Sep;132(3):451-5.

Kim J, Mayfield SP. Protein disulfide isomerase as a regulator of chloroplast translational activation. *Science*. 1997 Dec 12;278(5345):1954-7.

Kim PS, Arvan P. Endocrinopathies in the family of endoplasmic reticulum (ER) storage diseases: disorders of protein trafficking and the role of ER molecular chaperones. *Endocr Rev*. 1998 Apr;19(2):173-202.

Kivirikko KI, Myllyharju J. Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biol*. 16(7):357-68 (1998).

Klappa P, Freedman RB, Langenbuch M, Lan MS, Robinson GK, Ruddock LW. The pancreas-specific protein disulphide-isomerase PDIP interacts with a hydroxyaryl group in ligands. *Biochem J*. 2001 Mar 15;354(Pt 3):553-9.

Klappa P, Ruddock LW, Darby NJ, Freedman RB. The b' domain provides the principal peptide-binding site of protein disulfide isomerase but all domains contribute to binding of misfolded proteins. *EMBO J*. 1998 Feb 16;17(4):927-35.

Klappa P, Stromer T, Zimmermann R, Ruddock LW, Freedman RB. A pancreas-specific glycosylated protein disulphide-isomerase binds to misfolded proteins and peptides with an interaction inhibited by oestrogens. *Eur J Biochem*. 1998 May 15;254(1):63-9.

Knoblach B, Keller BO, Groenendyk J, Aldred S, Zheng J, Lemire BD, Li L, Michalak M. ERp19 and ERp46, new members of the thioredoxin family of endoplasmic reticulum proteins. *Mol Cell Proteomics*. 2003. (in press).

Koivunen P, Pirneskoski A, Karvonen P, Ljung J, Helaakoski T, Notbohm H, Kivirikko KI. The acidic C-terminal domain of protein disulfide isomerase is not critical for the enzyme subunit function or for the chaperone or disulfide isomerase activities of the polypeptide. *EMBO J.* 1999 Jan 4;18(1):65-74.

Kondrashov FA, Koonin EV. Evolution of alternative splicing: deletions, insertions and origin of functional parts of proteins from intron sequences. *Trends Genet.* 2003 Mar;19(3):115-9.

Konsolaki M, Schupbach T. windbeutel, a gene required for dorsoventral patterning in *Drosophila*, encodes a protein that has homologies to vertebrate proteins of the endoplasmic reticulum. *Genes Dev.* 1;12(1):120-31 (1998).

Kostova, Z. Wolf, D. H. For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection. *Embo J.* 22 (10): 2309-17 (2003).

Kramer B, Ferrari DM, Klappa P, Pohlmann N, Soling HD. Functional roles and efficiencies of the thioredoxin boxes of calcium-binding proteins 1 and 2 in protein folding. *Biochem J.* 2001 Jul 1;357(Pt 1):83-95.

Kriventseva EV, Koch I, Apweiler R, Vingron M, Bork P, Gelfand MS, Sunyaev S. Increase of functional diversity by alternative splicing. *Trends Genet.* 2003 Mar;19(3):124-8.

Kroning H, Kahne T, Ittenson A, Franke A, Ansorge S. Thiol-protein disulfide-oxidoreductase (protein disulfide isomerase): a new plasma membrane constituent of mature human B lymphocytes. *Scand J Immunol.* 1994 Apr;39(4):346-50.

Kuznetsov G, Bush KT, Zhang PL, Nigam SK. Perturbations in maturation of secretory proteins and their association with endoplasmic reticulum chaperones in a cell culture model for epithelial ischemia. *Proc Natl Acad Sci U S A.* 1996 Aug 6;93(16):8584-9.

Lahav J, Gofer-Dadosh N, Luboshitz J, Hess O, Shaklai M. Protein disulfide isomerase mediates integrin-dependent adhesion. *FEBS Lett.* 2000 Jun 16;475(2):89-92.

Lamberg A, Jauhainen M, Metso J, Ehnholm C, Shoulders C, Scott J, Pihlajaniemi T, Kivirikko KI. The role of protein disulphide isomerase in the microsomal triacylglycerol transfer protein does not reside in its isomerase activity. *Biochem J.* 1996 Apr 15;315 (Pt 2):533-6.

Langenbach KJ, Sottile J. Identification of protein-disulfide isomerase activity in fibronectin. *J Biol Chem.* 1999 Mar 12;274(11):7032-8.

LeMosy EK, Tan YQ, Hashimoto C. Activation of a protease cascade involved in patterning the *Drosophila* embryo. *Proc Natl Acad Sci U S A.* 2001 Apr 24;98(9):5055-60.

Liepinsh E, Baryshev M, Sharipo A, Ingelman-Sundberg M, Otting G, Mkrtchian S. Thioredoxin fold as homodimerization module in the putative chaperone ERp29: NMR structures of the domains and experimental model of the 51 kDa dimer. *Structure (Camb).* 2001 Jun;9(6):457-71.

Liour SS, Yu RK. Differentiation of radial glia-like cells from embryonic stem cells. *Glia.* 2003 Apr 15;42(2):109-17.

Li Y, Musacchio M, Finkelstein R. A homologue of the calcium-binding disulfide isomerase CaBP1 is expressed in the developing CNS of *Drosophila melanogaster*. *Dev Genet*. 1998;23(2):104-10.

Lumb RA, Bulleid NJ. Is protein disulfide isomerase a redox-dependent molecular chaperone? *EMBO J*. 2002 Dec 16;21(24):6763-70.

Ma Q, Guo C, Barnewitz K, Sheldrick GM, Soling HD, Uson I, Ferrari DM. Crystal structure and functional analysis of *drosophila* wind - a protein-disulfide isomerase-related protein. *J Biol Chem*. 2003 (in press)

MacLennan DH, Abu-Abed M, Kang C. Structure-function relationships in Ca^{2+} cycling proteins. *J Mol Cell Cardiol*. 2002 Aug;34(8):897-918.

MacLennan DH, Wong PT. Isolation of a calcium-sequestering protein from sarcoplasmic reticulum. *Proc Natl Acad Sci U S A*. 1971 Jun;68(6):1231-5.

Martin JL. Thioredoxin--a fold for all reasons. *Structure*. 1995 Mar 15;3(3):245-50.

Martin JL, Bardwell JC, Kuriyan J. Crystal structure of the DsbA protein required for disulphide bond formation in vivo. *Nature*. 1993 Sep 30;365(6445):464-8.

Matthias LJ, Hogg PJ. Redox Control on the Cell Surface: Implications for HIV-1 Entry. *Antioxid Redox Signal*. 2003 Feb;5(1):133-8.

Matthias LJ, Yam PT, Jiang XM, Vandegraaff N, Li P, Pountourios P, Donoghue N, Hogg PJ. Disulfide exchange in domain 2 of CD4 is required for entry of HIV-1. *Nat Immunol*. 2002 Aug;3(8):727-32.

McCarthy AA, Haebel PW, Torronen A, Rybin V, Baker EN, Metcalf P. Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. *Nat Struct Biol*. 2000 Mar;7(3):196-9.

Merry CL, Wilson VA. Role of heparan sulfate-2-O-sulfotransferase in the mouse. *Biochim Biophys Acta*. 2002 Dec 19;1573(3):319-27.

Mezghrani A, Fassio A, Benham A, Simmen T, Braakman I, Sitia R. Manipulation of oxidative protein folding and PDI redox state in mammalian cells. *EMBO J*. 2001 Nov 15;20(22):6288-96.

Missiakas D, Georgopoulos C, Raina S. Identification and characterization of the *Escherichia coli* gene *dsbB*, whose product is involved in the formation of disulfide bonds in vivo. *Proc Natl Acad Sci U S A*. 1993 Aug 1;90(15):7084-8.

Misson JP, Edwards MA, Yamamoto M, Caviness VS Jr. Mitotic cycling of radial glial cells of the fetal murine cerebral wall: a combined autoradiographic and immunohistochemical study. *Brain Res*. 1988 Feb 1;466(2):183-90.

Misson JP, Edwards MA, Yamamoto M, Caviness VS Jr. Identification of radial glial cells within the developing murine central nervous system: studies based upon a new

immunohistochemical marker. *Brain Res Dev Brain Res*. 1988 Nov 1;44(1):95-108. (*Brain Res Dev Brain Res*. 1988 Nov 1;44(1):95-108).

Mkrtchian S, Baryshev M, Matvijenko O, Sharipo A, Sandalova T, Schneider G, Ingelman-Sundberg M, Mkrtchiana S. Oligomerization properties of ERp29, an endoplasmic reticulum stress protein. *FEBS Lett*. 1998 Jul 24;431(3):322-6.

Molinari M, Helenius A. Glycoproteins form mixed disulphides with oxidoreductases during folding in living cells. *Nature*. 1999 Nov 4;402(6757):90-3

Monnat J, Hacker U, Geissler H, Rauchenberger R, Neuhaus EM, Maniak M, Soldati T. Dictyostelium discoideum protein disulfide isomerase, an endoplasmic reticulum resident enzyme lacking a KDEL-type retrieval signal. *FEBS Lett*. 1997 Dec 1;418(3):357-62.

Monnat J, Neuhaus EM, Pop MS, Ferrari DM, Kramer B, Soldati T. Identification of a novel saturable endoplasmic reticulum localization mechanism mediated by the C-terminus of a Dictyostelium protein disulfide isomerase. *Mol Biol Cell*. 2000 Oct;11(10):3469-84.

Neil J. Bulleid. Protein disulphide isomerase. *Curr Biol*. 2003 May 13;13(10):R380.

Neuman-Silberberg FS, Schupbach T. The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGF alpha-like protein. *Cell*. 1993 Oct 8;75(1):165-74.

Nicchitta CV, Blobel G. Luminal proteins of the mammalian endoplasmic reticulum are required to complete protein translocation. *Cell*. 1993 Jun 4;73(5):989-98.

Nigam SK, Goldberg AL, Ho S, Rohde MF, Bush KT, Sherman MYu. A set of endoplasmic reticulum proteins possessing properties of molecular chaperones includes Ca(2+)-binding proteins and members of the thioredoxin superfamily. *J Biol Chem*. 1994 Jan 21;269(3):1744-9.

Nilson LA, Schupbach T. Localized requirements for windbeutel and pipe reveal a dorsoventral prepattern within the follicular epithelium of the Drosophila ovary. *Cell*. 1998 Apr 17;93(2):253-62.

Noiva, R. Protein disulfide isomerase: the multifunctional redox chaperone of the endoplasmic reticulum. *Semin Cell Dev Biol*. 10, 481-93 (1999).

Nordstrand K, slund F, Holmgren A, Otting G, Berndt KD. NMR structure of Escherichia coli glutaredoxin 3-glutathione mixed disulfide complex: implications for the enzymatic mechanism. *J Mol Biol*. 1999 Feb 19;286(2):541-52.

Oliver JD, Roderick HL, Llewellyn DH, High S. ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin. *Mol Biol Cell*. 1999 Aug;10(8):2573-82.

Oliver JD, van der Wal FJ, Bulleid NJ, High S. Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science*. 1997 Jan 3;275(5296):86-8.

O'Neill S, Robinson A, Deering A, Ryan M, Fitzgerald DJ, Moran N. The platelet integrin α IIb β 3 has an endogenous thiol isomerase activity. *J Biol Chem*. 2000 Nov 24;275(47):36984-90.

Parnavelas JG, Nadarajah B. Radial glial cells. are they really glia? *Neuron*. 2001 Sep 27;31(6):881-4.

Paschen W, Gissel C, Linden T, Doutheil J. ERp72 expression activated by transient cerebral ischemia or disturbance of neuronal endoplasmic reticulum calcium stores. *Metab Brain Dis*. 1998 Mar;13(1):55-68.

Pihlajaniemi T, Helaakoski T, Tasanen K, Myllyla R, Huhtala ML, Koivu J, Kivirikko KI. Molecular cloning of the beta-subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. *EMBO J*. 1987 Mar;6(3):643-9.

Pirneskoski A, Ruddock LW, Klappa P, Freedman RB, Kivirikko KI, Koivunen P. Domains b' and a' of protein disulfide isomerase fulfill the minimum requirement for function as a subunit of prolyl 4-hydroxylase. The N-terminal domains a and b enhances this function and can be substituted in part by those of ERp57. *J Biol Chem*. 2001 Apr 6;276(14):11287-93.

Puig A, Gilbert HF. Protein disulfide isomerase exhibits chaperone and anti-chaperone activity in the oxidative refolding of lysozyme. *J Biol Chem*. 1994 Mar 11;269(10):7764-71.

Qin J, Clore GM, Kennedy WM, Huth JR, Gronenborn AM. Solution structure of human thioredoxin in a mixed disulfide intermediate complex with its target peptide from the transcription factor NF kappa B. *Structure*. 1995 Mar 15;3(3):289-97.

Ren B, Tibbelin G, de Pascale D, Rossi M, Bartolucci S, Ladenstein R. A protein disulfide oxidoreductase from the archaeon *Pyrococcus furiosus* contains two thioredoxin fold units. *Nat Struct Biol*. 1998 Jul;5(7):602-11.

Rietsch A, Belin D, Martin N, Beckwith J. An in vivo pathway for disulfide bond isomerization in *Escherichia coli*. *Proc Natl Acad Sci U S A*. 1996 Nov 12;93(23):13048-53.

Ruoppolo M, Orru S, Talamo F, Ljung J, Pirneskoski A, Kivirikko KI, Marino G, Koivunen P. Mutations in domain a' of protein disulfide isomerase affect the folding pathway of bovine pancreatic ribonuclease A. *Protein Sci*. 2003 May;12(5):939-52.

Rupp K, Birnbach U, Lundstrom J, Van PN, Soling HD. Effects of CaBP2, the rat analog of ERp72, and of CaBP1 on the refolding of denatured reduced proteins. Comparison with protein disulfide isomerase. *J Biol Chem*. 1994 Jan 28;269(4):2501-7.

Ryser HJ, Levy EM, Mandel R, DiSciullo GJ. Inhibition of human immunodeficiency virus infection by agents that interfere with thiol-disulfide interchange upon virus-receptor interaction. *Proc Natl Acad Sci U S A*. 1994 May 10;91(10):4559-63.

Safran M, Farwell AP, Leonard JL. Thyroid hormone-dependent redistribution of the 55-kilodalton monomer of protein disulfide isomerase in cultured glial cells. *Endocrinology*. 1992 Nov;131(5):2413-8.

Safran M, Leonard JL. Characterization of a N-bromoacetyl-L-thyroxine affinity-labeled 55-kilodalton protein as protein disulfide isomerase in cultured glial cells. *Endocrinology*. 1991 Oct;129(4):2011-6.

Sargsyan E, Baryshev M, Szekely L, Sharipo A, Mkrtchian S. Identification of ERp29, an endoplasmic reticulum lumenal protein, as a new member of the thyroglobulin folding complex. *J Biol Chem*. 2002 May 10;277(19):17009-15.

Sarnat HB. Histochemistry and immunocytochemistry of the developing ependyma and choroid plexus. *Microsc Res Tech*. 1998 Apr 1;41(1):14-28.

Selleck SB. Proteoglycans and pattern formation: sugar biochemistry meets developmental genetics. *Trends Genet*. 2000 May;16(5):206-12.

Sen J, Goltz JS, Konsolaki M, Schupbach T, Stein D. Windbeutel is required for function and correct subcellular localization of the *Drosophila* patterning protein Pipe. *Development*. 127(24):5541-50 (2000).

Sen J, Goltz JS, Stevens L, Stein D. Spatially restricted expression of pipe in the *Drosophila* egg chamber defines embryonic dorsal-ventral polarity. *Cell*. 1998 Nov 13;95(4):471-81.

Sergeev P, Streit A, Heller A, Steinmann-Zwicky M. The *Drosophila* dorsoventral determinant PIPE contains ten copies of a variable domain homologous to mammalian heparan sulfate 2-sulfotransferase. *Dev Dyn*. 2001 Feb;220(2):122-32.

Sevier CS, Cuozzo J, Vala A, Åslund F and Kaiser CA. A flavoprotein oxidase defines a new endoplasmic reticulum pathway for biosynthetic disulphide bond formation. *Nat Cell Biol*. 2001 Oct;3(10):874-82.

Sevier CS, Kaiser CA. Formation and transfer of disulphide bonds in living cells. *Nat Rev Mol Cell Biol*. 2002 Nov;3(11):836-47.

Shnyder SD, Hubbard MJ. ERp29 is a ubiquitous resident of the endoplasmic reticulum with a distinct role in secretory protein production. *J Histochem Cytochem*. 2002 Apr;50(4):557-66.

Song JL, Wang CC. Chaperone-like activity of protein disulfide-isomerase in the refolding of rhodanese. *Eur J Biochem*. 1995 Jul 15;231(2):312-6.

Speake T, Whitwell C, Kajita H, Majid A, Brown PD. Mechanisms of CSF secretion by the choroid plexus. *Microsc Res Tech*. 2001 Jan 1;52(1):49-59.

Sullivan DC, Huminiecki L, Moore JW, Boyle JJ, Poulson R, Creamer D, Barker J, Bicknell R. EndoPDI, a novel protein disulphide isomerase-like protein that is preferentially expressed in endothelial cells acts as a stress survival factor. *J Biol Chem*. 2003 (in press).

Tachikawa H, Funahashi W, Takeuchi Y, Nakanishi H, Nishihara R, Katoh S, Gao XD, Mizunaga T, Fujimoto D. Overproduction of Mpd2p suppresses the lethality of protein disulfide isomerase depletion in a CXXC sequence dependent manner. *Biochem Biophys Res Commun*. 1997 Oct 29;239(3):710-4.

Terada K, Manchikalapudi P, Noiva R, Jauregui HO, Stockert RJ, Schilsky ML. Secretion, surface localization, turnover, and steady state expression of protein disulfide isomerase in rat hepatocytes. *J Biol Chem*. 1995 Sep 1;270(35):20410-6.

Tsai B, Rapoport TA. Unfolded cholera toxin is transferred to the ER membrane and released from protein disulfide isomerase upon oxidation by Ero1. *J Cell Biol*. 2002 Oct 28;159(2):207-16.

Tsai B, Rodighiero C, Lencer WI, Rapoport TA. Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. *Cell*. 2001 Mar 23;104(6):937-48.

Tsai B, Ye Y, Rapoport TA. Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat Rev Mol Cell Biol*. 2002 Apr;3(4):246-55.

Tu BP, Ho-Schleyer SC, Travers KJ, Weissman JS. Biochemical basis of oxidative protein folding in the endoplasmic reticulum. *Science*. 2000 Nov 24;290(5496):1571-4.

Tu BP, Weissman JS. The FAD- and O₂-dependent reaction cycle of Ero1-mediated oxidative protein folding in the endoplasmic reticulum. *Mol Cell*. 2002 Nov;10(5):983-94.

Van der Wal FJ, Oliver JD, High S. The transient association of ERp57 with N-glycosylated proteins is regulated by glucose trimming. *Eur J Biochem*. 1998 Aug 15;256(1):51-9.

Van Eeden F, St Johnston D. The polarisation of the anterior-posterior and dorsal-ventral axes during *Drosophila* oogenesis. *Curr Opin Genet Dev*. 1999 Aug;9(4):396-404.

van Endert PM. Genes regulating MHC class I processing of antigen. *Curr Opin Immunol*. 1999 Feb;11(1):82-8.

Van PN, Peter F, Soling HD. Four intracisternal calcium-binding glycoproteins from rat liver microsomes with high affinity for calcium. No indication for calsequestrin-like proteins in inositol 1,4,5-trisphosphate-sensitive calcium sequestering rat liver vesicles. *J Biol Chem*. 1989 Oct 15;264(29):17494-501.

Van PN, Rupp K, Lampen A, Soling HD. CaBP2 is a rat homolog of ERp72 with protein disulfide isomerase activity. *Eur J Biochem*. 1993 Apr 15;213(2):789-95.

Veijola J, Pihlajaniemi T, Kivirikko KI. Co-expression of the alpha subunit of human prolyl 4-hydroxylase with BiP polypeptide in insect cells leads to the formation of soluble and insoluble complexes. Soluble alpha-subunit-BiP complexes have no prolyl 4-hydroxylase activity. *Biochem J*. 1996 Apr 15;315 (Pt 2):613-8.

Volkmer J, Guth S, Nastainczyk W, Knippel P, Klappa P, Gnau V, Zimmermann R. Pancreas specific protein disulfide isomerase, PDIP, is in transient contact with secretory proteins during late stages of translocation. *FEBS Lett*. 1997 Apr 14;406(3):291-5.

Wadsworth SC, Vincent WS 3rd, Bilodeau-Wentworth D. A *Drosophila* genomic sequence with homology to human epidermal growth factor receptor. *Nature*. 1985 Mar 14-20;314(6007):178-80.

Wang CC. Protein disulfide isomerase assists protein folding as both an isomerase and a chaperone. *Ann N Y Acad Sci.* 13;864:9-13 (1998).

Wang CC. Isomerase and chaperone activities of protein disulfide isomerase are both required for its function as a foldase. *Biochemistry (Mosc).* 63(4):407-12 (1998);.

Wang CC, Tsou CL. Enzymes as chaperones and chaperones as enzymes. *FEBS Lett.* 3;425(3):382-4 (1998).

Wang S, Trumble WR, Liao H, Wesson CR, Dunker AK, Kang CH. Crystal structure of calsequestrin from rabbit skeletal muscle sarcoplasmic reticulum. *Nat Struct Biol.* 1998 Jun;5(6):476-83.

Wetterau JR, Combs KA, Spinner SN, Joiner BJ. Protein disulfide isomerase is a component of the microsomal triglyceride transfer protein complex. *J Biol Chem.* 1990 Jun 15;265(17):9801-7.

Willnow TE, Rohlmann A, Horton J, Otani H, Braun JR, Hammer RE, Herz J. RAP, a specialized chaperone, prevents ligand-induced ER retention and degradation of LDL receptor-related endocytic receptors. *EMBO J.* 1996 Jun 3;15(11):2632-9.

Woycechowsky KJ, Raines RT. The CXC motif: a functional mimic of protein disulfide isomerase. *Biochemistry.* 2003 May 13;42(18):5387-94.

Yao Y, Zhou Y, Wang C. Both the isomerase and chaperone activities of protein disulfide isomerase are required for the reactivation of reduced and denatured acidic phospholipase A2. *EMBO J.* 1997 Feb 3;16(3):651-8.

Yokoi T, Nagayama S, Kajiwarra R, Kawaguchi Y, Horiuchi R, Kamataki T. Identification of protein disulfide isomerase and calreticulin as autoimmune antigens in LEC strain of rats. *Biochim Biophys Acta.* 1993 Nov 28;1158(3):339-44.

Zapun A, Darby NJ, Tessier DC, Michalak M, Bergeron JJ, Thomas DY. Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57. *J Biol Chem.* 1998 Mar 13;273(11):6009-12.

Zapun A, Missiakas D, Raina S, Creighton TE. Structural and functional characterization of DsbC, a protein involved in disulfide bond formation in *Escherichia coli*. *Biochemistry.* 1995 Apr 18;34(15):5075-89.

Zhang SC. Defining glial cells during CNS development. *Nat Rev Neurosci.* 2001 Nov;2(11):840-3.

Publication:

Ma Q*, Guo C*, Barnewitz K, Sheldrick GM, Soling HD, Uson I, Ferrari DM. Crystal structure and functional analysis of drosophila Wind, protein-disulfide isomerase-related protein. J Biol Chem. 2003 (in press)

* equally contributed to this paper

Curriculum vitae

Name: Chaoshe Guo
Date of Birth: 22th April 1971
Place of Birth: Henan Province, P.R.China
Marital status: Married
Nationality: P.R.China

Education and research career:

1978-1983	Primary school, Henan, China
1983-1986	Middel school, Henan, China
1986-1989	High school, Henan, China
1989-1994	Department of Clinical Medicine, Henan Medical University, Zhengzhou, China for bachelor degree, including one year as intern in Affiliated Hospital of Henan Medical University.
1994	Bachelor of Medicine
1994-1996	Research Assistant in the Department of Pharmacology, Hennan Medical University, China
1996-1999	Peking Union Medical College (Chinese Academy of Medical Science)
1999	Master of Science, in Medical Molecular Biology
Jan, 2000-Jul, 2000	Researcher in the Department of Neurobiology, ZMBE, University of Münster, Münster, Germany
Jul, 2000-2003	Dept. of Neurobiology, Max Planck Institute of Biophysical Chemistry, Göttingen, Germany for Ph.D degree.