# Structure-function analysis of the mammalian peroxidasin protein

Doctoral theses

**Dr. Enikő Marina Lázár**

Semmelweis University
Molecular Medicine PhD School

Supervisor: Dr. Miklós Geiszt, D.Sc., professor

Reviewers: Dr. Tamás Kardon, PhD, associate professor
Dr. Gergely Szakács, PhD, senior research fellow

Chairman of the Final Exam Committee: Dr. Balázs Sarkadi, member of the HAS, professor
Members of the Final Exam Committee: Dr. László Cervenak, PhD, senior research fellow
Dr. Péter Nagy, PhD, academic head of department

Budapest
2016

# INTRODUCTION

Reactive oxygen species (ROS) have long been considered barely as toxic by-products of cellular oxidative metabolism. However, intensive research in the past decades has revealed that these oxygen-containing intermediates play a central role in a number of physiological processes, including the development of extracellular matrix (ECM) structure. One of the most well-studied ROS molecules is hydrogen peroxide ($H_2O_2$) which often exerts its effects via peroxidase enzymes. Mammalian members of the animal heme peroxidase family include myeloperoxidase (MPO), eosinophil peroxidase (EPX), lactoperoxidase (LPO) and thyroid peroxidase (TPO), as well as cyclooxygenase 1 and 2 (COX1, COX2), all of which display a high degree of sequence similarity and identical catalytic cycle. My doctoral thesis focuses on the most recently identified members of the animal heme peroxidase family, peroxidasin (PXDN) and its loss-of-function mutant homologue, peroxidasin-like protein (PXDNL).

PXDN possesses certain unique structural features as it contains elements characteristic of matrix proteins, in addition to its catalytic peroxidase domain. Its N-terminal secretion signal sequence is followed by six leucine-rich repeats (LRR) and four immunoglobulin C2-like domains (Ig C2). C-terminally PXDN contains an α-helical region and a von Willebrand factor type C domain (vWF C). While PXDN was first described in *Drosophila melanogaster*, later it was found to be ubiquitously expressed throughout the animal kingdom. *Drosophila* PXDN was shown to be secreted by hemocytes to the extracellular space in a homotrimeric form but the mechanism of oligomerization and the protein regions involved in the process have been unknown so far. While human PXDN was first described as a melanoma-associated gene product, later on it was found to be present in a number of cell lines and a wide range of tissues. Our group was the first to identify and characterize peroxidasin-like protein (PXDNL), the heart-specific homologue of PXDN. Unlike PXDN, PXDNL does not possess peroxidase activity due to mutations affecting two conserved amino acids essential for heme binding in other heme peroxidase enzymes.

Available information has been scarce regarding the *in vivo* function of the two proteins but PXDN has long been suspected to act as an enzyme specialized for the oxidative modification of the ECM. The importance of PXDN in the development of certain ECM structures has been established in various developmental stages of lower

species (*Caenorhabditis elegans*, *Drosophila melanogaster*, *Xenopus tropicalis*) and our group has previously described the accumulation of the protein in the extracellular space of the fibrotic mouse kidney following ureteral obstruction. A number of studies have identified mutations in the *pxdn* gene as the causative factor behind complex eye-developmental and musculoskeletal disturbances in human patients. The heterogeneous eye manifestations termed anterior segment dysgenesis (ASD) were studied in detail using PXDN knockout mouse strain. While these results confirmed the importance of PXDN in ECM homeostasis, the molecular mechanisms influencing the integrity of this complex protein structure remained largely unknown.

A recent breakthrough in the field was the identification of the first physiological enzymatic function of PXDN which occurs in basal membranes. These specialized forms of ECM are flat structures found below endothelial and epithelial cells as well as around muscle cells, adipocytes and peripheral nerves. Basal membranes provide structural support and also convey regulatory signals from the ECM towards the cells connected to them. The most prevalent component of basal membranes is collagen IV which undergoes a multistep oligomerization process during ECM synthesis. All six isoforms of collagen IV α-chains contain an N-terminal 7S region, a long collagenous domain and a C-terminal NC1-domain. The functional units of collagen IV are the protomers which are formed by three α-chains and are stabilized and connected to each other by several covalent and non-covalent bonds. The NC1 domains of two adjacent protomers create a hexameric structure and these dimeric protomers are further organized into tetrameric complexes via the 7S domains. These interactions provide the structural basis for the assembly of the two-dimensional collagen IV network.

PXDN was recently described to create a covalent bond between the NC1 domains of two adjacent collagen IV protomers in the ECM, thus contributing to the synthesis of fully functional basal membranes. In this reaction PXDN utilizes hypobromous acid (HOBr) generated from $H_2O_2$ and bromide ions, to crosslink the sidechains of methionine and the lysine- or hydroxylysine residues present in the 93. an 211. positions of the collagen IV sequence with a covalent sulfilimine bond. This catalytic process displays two unique features: this was the first report of the presence of sulfilimine bond in a biological system and it also provides the first evidence for the essential role of bromine in living organisms.

These results were primarily based on *in vitro* experiments investigating the effects of pharmacological inhibition of peroxidase activity in the PFHR-9 embryonic carcinoma cell line as well as the biochemical analysis of heterologously expressed recombinant PXDN. The *in vivo* significance of NC1 crosslinking by PXDN has only been confirmed in *Drosophila melanogaster* and *Danio rerio* so far. While the high level of evolutionary conservation of all members in this process suggests a similar mechanism in other species, there has been no mammalian data supporting this hypothesis so far. It is also unclear how distinct structural domains of PXDN contribute to its oligomerization and known functions including collagen IV crosslinking. Furthermore, in the case of PXDN knockout animals it is still uncertain which phenotypic effects are due to the loss of collagen IV crosslinking reaction, or to the loss of other oxidation processes of yet unknown substrates in the ECM, or possibly to the lack of the protein's structural role in the matrix.

## OBJECTIVES

The goal of my doctoral work was to develop a better understanding of the relationship between the special structure and the physiological function of PXDN. Initially, I attempted to study the interaction between endogenous mammalian PXDN and collagen IV in an experimental setup which provided more physiological conditions compared to the systems previously described in the literature. We planned to investigate the role of specific regions of PXDN sequence in the establishment of the protein's quaternary structure as well as its enzymatic activity, transport and localization by a structure-function analysis using recombinant protein constructs. Based on the results we hoped to draw additional conclusions regarding the structure and function of the highly homologous PXDNL.

In accordance with these considerations we have laid out the following objectives:

1. the analysis of endogenous mammalian PXDN's role in the crosslinking of collagen IV in primary cell cultures isolated from living tissue,
2. the investigation of the quaternary structure of mammalian PXDN and the mechanism of its potential oligomerization,
3. the examination of structural requirements necessary for the known physiological functions of PXDN, including peroxidase activity, secretion and adhesion,
4. the study of similarities in both sequence and subcellular localization between PXDN and PXDNL.

## METHODS

*Cloning of recombinant human PXDN and PXDNL constructs:* Based on the wild-type human PXDN and PXDNL plasmids, we created a number of V5-tagged recombinant protein constructs modified by point mutations and domain deletions. Truncated constructs were generated by inserting particular regions of PXDN coding sequence into pcDNA3.1/V5-His-TOPO vector using TA cloning strategy. Targeted mutations were introduced into the full-length sequence with double-primer PCR method using complementary mutagenic oligonucleotide pairs.

*Generation and breeding of PXDN knockout mice:* SAGE Laboratories have generated a PXDN knockout C57BL/6 mouse strain using zinc-finger endonuclease (ZFN) technology. 13 days after the mating of heterozygous breeding pairs, pregnant females were sacrificed. Embryos removed from the uterus were used for the isolation of mouse embryonic fibroblasts (MEFs).

*Cell culture experiments:* Recombinant PXDN and PXDNL constructs were expressed in HEK293 freestyle (FS) and Cos7 cells following transfection with Lipofectamine LTX and Plus reagents. By geneticin-selection of HEK293FS cells transfected with pcDNA3.1/hsPXDN-V5-His plasmid we created PXDN-V5-expressing stable cell lines. Endogenous human PXDN was studied in human umbilical vein endothelial cells (HUVECs) where we used the general peroxidase inhibitor phloroglucinol to test the protein's collagen IV crosslinking activity. The effects resulting from the total absence of PXDN caused by genomic alterations was investigated in wild-type and PXDN-deficient MEFs. For rescue experiments recombinant PXDN constructs were introduced into knockout MEFs via electroporation with Neon transfection system.

*Investigation of PXDN oligomerization:* Total cell lysates from HUVEC, MEF and transiently transfected HEK293FS cells were tested in Western blots following treatment with the thiol alkylating agent N-ethylmaleimide. The oligomerization of PXDN was evaluated in non-reduced samples based on the appearance of PXDN- or V5-specific signals corresponding to the molecular weight of the monomeric, dimeric and trimeric PXDN forms.

*Analysis of collagen IV profile and NC1 crosslinking:* The analysed cell culture was suspended in hypotonic lysis buffer and treated with collagenase enzyme to release the NC1 domains from the insoluble, matrix-bound collagen IV. Monomeric and dimeric NC1 domains in the digested samples were separated using SDS-PAGE and identified at 25 and 55 kDa using isoform-specific antibodies. Collagen IV crosslinking activity of the recombinant PXDN constructs was tested in rescue experiments following transfection of PXDN-deficient MEFs.

*Analysis of PXDN secretion:* The secretion of different recombinant PXDN constructs was compared based on the PXDN content of the cell culture medium and total cell lysates of transiently transfected Cos7 cells. We used α-V5 antibody to detect the recombinant proteins in Western blot experiments, while α-β-actin antibody was used as loading control in the case of cell lysates and also to confirm the lack of cellular contamination in the medium samples.

*Western blotting and immunoprecipitation:* Reduced and non-reduced protein samples were analysed in Western blot experiments using polyclonal α-hsPXDN or α-mmPXDN antibodies developed in-house or monoclonal α-NC1 collagen IV, α-V5 or α-β-actin antibodies. Anti-V5 agarose affinity gel was used to isolate recombinant PXDN-V5 oligomers from the stable HEK293FS-derived cell line. Protein complexes formed by endogenous PXDN in HUVECs were precipitated using α-hsPXDN antibody and captured with protein G agarose beads.

*Mass spectrometric analysis:* The composition of the PXDN-containing protein complexes was analysed by our collaborators based on the LC-MS/MS spectra gained from the triptic digests of the samples. The results were compared to the latest non-redundant protein database available at NCBI using a Mascot server.

*Immunofluorescent staining and confocal microscopy:* The localisation of PXDN and PXDNL constructs was analysed in PFA-fixed samples of transiently transfected Cos7 and PXDN-deficient MEFs. With or without prior permeabilization, cells were

incubated with α-hsPXDN or α-V5 primary antibodies and subsequently stained using Alexa Fluor 488 or 568 secondary antibodies. Microscopic pictures of the samples were taken using an LSM 710 confocal microscope with a 63x objective.

*Measurement of peroxidase activity:* Cell lysates from Cos7 cells transiently transfected with recombinant PXDN constructs were analysed for *in vitro* peroxidase activity in Amplex Red peroxidase assay. Fluorescence of resorufin produced in the reaction was detected at 590 nm using a POLARstar OPTIMA 96-well plate reader.

*Multiple sequence alignment:* The Clustal Omega software was used to compare sequences of human heme peroxidases, PXDN from various species as well as human PXDN and PXDNL.

*Statistical analysis:* All graphs presented in the thesis are representative of results from at least 3 independent experiments. One-way ANOVA with Sidak's post-hoc test was used to compare the enzymatic activity of various constructs in the Amplex Red peroxidase assay. All statistical analyses were done using the GraphPad PRISM 5 software. Significance was assumed at $p<0.05$.

## RESULTS

Crosslinking ability of endogenous mammalian PXDN was analysed using human and mouse primary cell cultures. NC1 crosslinking activity of human PXDN was studied in human umbilical vein endothelial cells (HUVECs) using the general peroxidase-inhibitor phloroglucinol (PhG). HUVECs express α1 and α2 isoforms of collagen IV in high, while α5 isoform only in smaller quantities. Western blots confirmed that PhG treatment of HUVECs had no effect on the expression of endogenous PXDN, while drastically decreased the amount of dimeric collagen IV NC1 domains detected in collagenase-digested cell lysates. We confirmed these pharmacological results with a genetic approach using wild-type and PXDN knockout mouse embryonic fibroblasts (MEFs). While Western blots showed high levels of dimeric α1 and α2 NC1 domains in samples of wild-type cells, these crosslinked forms were completely missing from PXDN-deficient cell lysates. Our results suggest that enzymatic activity of PXDN is essential for crosslinking of collagen IV NC1 domains in primary mammalian cells *in vitro*.

The quaternary structure of mammalian PXDN was analysed in HUVECs as well as in a stable HEK293FS cell line expressing recombinant human PXDN-V5. In Western blots of non-reduced cell lysates three PXDN-containing bands were detected at 160, 300 and 500 kDa with the highest band being the dominant one in both samples. These signals disappeared upon reduction with β-mercaptoethanol and only a single band could be detected at 165 kDa corresponding to the molecular weight of monomeric PXDN. In order to analyze their composition, PXDN-containing complexes were isolated from the cell lysates. Coomassie stained protein bands at 500, 300 and 160 kDa were detected in samples immunoprecipitated with α-V5 antibody from the stable HEK cell line. Proteomic analysis identified PXDN as the only component of all three protein bands. We also purified the 500 kDa PXDN-positive protein complex from HUVECs and got the same result in the mass spectrometric analysis of the sample. Thus we concluded that mammalian PXDN, similarly to the *Drosophila* protein, forms a homotrimer which is sensitive to reducing agents and the observed complexes therefore correspond to the various stages of the oligomerization process. In subsequent

experiments the mechanism of homotrimerization was studied using recombinant PXDN constructs.

In order to investigate the connection between the oligomer formation and endogenous peroxidase activity of PXDN, a loss-of-function mutant was created from the wild-type protein by site-directed mutagenesis. Two amino acid residues essential for catalytic activity were replaced by the corresponding amino acids of the PXDNL sequence, thus creating the Q823W, D826E construct. *In vitro* peroxidase activity of the mutant was tested in Amplex Red peroxidase assay with lysates of transiently transfected Cos7 cells. The introduced mutations resulted in the loss of peroxidase but did not affect the homotrimerization process of the protein in transiently transfected HEK293FS cells. We can thus conclude that its peroxidase activity is not necessary for the formation of trimeric quaternary structure of PXDN.

Subsequently we attempted to identify the amino acid regions in PXDN responsible for the assembly and stabilization of the trimeric form. Our first step was to create a set of domain-deficient PXDN construct and investigate if they retain their ability to form oligomers in transiently transfected HEK293FS cells. With Western blot. analysis of non-reduced cell lysates we showed that neither the lack of vWF C domain nor the lack of LRR and Ig C2 domains prevented oligomerization of PXDN as long as the secretion signal sequence was intact at the N-terminus of the protein. Based on these results we concluded that the association between the three subunits of the homotrimer is mediated by the region between the 696. and 1412. amino acids of the human PXDN which spans from the beginning of the peroxidase domain to the end of the α-helical region.

The sensitivity of the PXDN homotrimer to reducing agents suggested that disulphide bonds might be responsible for the stabilization of the protein complex. By comparing the sequences of human PXDN to other heme-peroxidases we identified four cysteine residues in the 736., 1315., 1316., and 1319. positions which presumably do not participate in the formation of intramolecular disulphide bridges and thus could potentially contribute to intermolecular bonds stabilizing the homotrimeric complex. By site-directed mutagenesis of these amino acid we created several cysteine-serine mutant constructs which were expressed in HEK293FS cells. Western blots of non-reduced cell lysates showed that neither the single, nor the double mutation of the 1316. and 1319.

cysteines altered the trimerization pattern of PXDN. In contrast, both the C736S and the C1315S single mutant constructs displayed strong signals at the monomeric and dimeric heights, while the highest molecular weight complex disappeared. The simultaneous mutation of both of these cysteines completely abolished the oligomerization process. These results suggest that the 736. and the 1315. cysteines participate in the formation of the homotrimer-stabilizing disulphide bonds. Multiple sequence alignment between species from various levels of evolutionary development (*Drosophila melanogaster*, *Caenorhabditis elegans*, *Xenopus tropicalis*, *Mus musculus*) confirmed the conservation of these cysteine residues. This suggests that the mechanism of trimer formation might also be conserved between the studied species and involve the cysteine residues corresponding to the 736. and 1315. positions in the human sequence.

We have attempted to identify the structural requirements of the known functions of PXDN in a structure-function analysis performed with the recombinant protein constructs introduced previously. Amino acid regions of PXDN necessary for its secretion were determined by Western blot analysis of the cell culture medium of Cos7 cells expressing various recombinant PXDN constructs. Regardless of any other structural characteristics of the tested PXDN forms, they appeared in the supernatant of the cells if the first 26 amino acids of the original sequence were preserved at the N-terminus of the protein. According to our results, the presence of the N-terminal signal sequence appeared to be both necessary and sufficient for the secretion of PXDN.

To identify the structural characteristics affecting PXDN's *in vitro* peroxidase activity, we expressed several modified PXDN constructs in Cos7 cells and compared the peroxidase activity of the cell lysates in Amplex Red peroxidase assay. The wild-type PXDN showed significantly higher enzymatic activity compared either to negative controls or to the loss-of-function mutant protein. Neither mutations of the 736. and 1315. cysteines resulting in monomeric PXDN, nor the lack of vWF C domain affected the peroxidase activity of PXDN. However, the construct lacking the LRR and Ig C2 domains displayed no enzymatic activity, presumably due to incorrect folding of the protein caused by the drastic N-terminal truncation.

Afterwards we investigated the structural prerequisites for the only known physiological function of PXDN, collagen IV crosslinking. For this aim PXDN knockout MEFs proved to be valuable tools which we used in rescue experiments. In

this experimental setup wild-type and mutant PXDN constructs were reintroduced into the cells, and then we examined the extent to which they were able to restitute the missing IV collagen crosslinking function of the cells. Wild-type and vWF C domain-deficient constructs could successfully catalyse covalent crosslinking of the NC1 domains, while the constructs lacking peroxidase activity showed no crosslinking potential. The monomeric C736S, C1315S construct retained its ability to form NC1 crosslinks but only to a significantly smaller degree compared to the wild-type protein.

In order to investigate the possible reasons behind the different crosslinking efficiency of the wild-type and the monomeric mutant forms, we studied the subcellular localization of heterologously expressed PXDN constructs in immunostaining experiments. In permeabilized samples of transiently transfected Cos7 cells all tested PXDN constructs were detected in the endoplasmic reticulum. Beyond that, we identified PXDN-positive granule-like structures along the plasma membrane in cells expressing the wild-type, the loss-of-function mutant and the vWF C domain-deficient protein, but not the monomeric PXDN form. Immunostaining was repeated without prior permeabilization of the cells in order to separate signals originating from the intra- and extracellular space. Unlike the protein fraction localized to the endoplasmic reticulum, recombinant PXDN in cell surface spots remained detectable in such conditions which confirmed its extracellular localization. No PXDN-specific signal could be observed in non-permeabilized cells expressing the monomeric protein. The disparate localization of the wild-type and the monomeric PXDN was also confirmed in PXDN knockout MEFs transfected with these constructs.

Based on our results we have developed a comprehensive model of the mechanism of PXDN oligomerization and its collagen IV crosslinking function. PXDN homotrimers are formed within the endoplasmic reticulum via intermolecular disulphide bonds between cysteine residues located in the 736. and 1315. positions of the protein sequence. Upon secretion to the extracellular space PXDN trimers form adherent structures on the cell surface which could presumably constitute ‘hot spots’ for the collagen IV crosslinking reaction. This model could account for the decreased crosslinking efficacy of the monomeric PXDN construct which failed to appear in the cell surface localization.

In the final part of my thesis I discuss the results of the sequence analysis and localization studies performed with the recombinant form of the PXDN homologue PXDNL. In his doctoral thesis my colleague, Dr. Zalán Péterfi described that under non-reducing conditions PXDNL participates in the formation of high molecular weight complexes similar to PXDN, which might reflect a possible oligomerization process. Sequence alignment between the two proteins showed that 736. and 1315. cysteines, which proved to be critical for PXDN trimerization, are conserved in the 722. and 1298. positions of PXDNL sequence, and are surrounded by highly similar amino acid regions as well. This suggests that PXDNL, following the same mechanism as PXDN, might assemble into a homotrimers stabilized by disulphide bonds between these cysteine residues.

While PXDNL does not possess enzymatic activity like PXDN, its matrix protein-like domains allow it to fulfil a potential adhesive role in the ECM. Cell surface adhesion of PXDNL was investigated by immunostaining of Cos7 cells expressing the recombinant V5-tagged PXDNL construct. To distinguish between extra- and intracellular staining, we compared cells processed with or without permeabilization prior to staining. In the permeabilized samples the recombinant protein appeared in dual localization including the endoplasmic reticulum and punctate structures along the cell membrane, while we could only detect the membrane-bound signal in non-permeabilized samples.

## CONCLUSIONS

1. The enzymatic activity of endogenous mammalian PXDN is essential for the covalent crosslinking of collagen IV protomers' NC1 domains in the basement membranes of primary mammalian cells.
2. PXDN assembles into a homotrimeric complex through an oligomerization process independent from its endogenous peroxidase activity. The PXDN homotrimer is stabilized by intermolecular disulphide bonds that are formed between highly conserved cysteine residues located inside and right after the C-terminal end of the peroxidase domain. In the human PXDN these cysteines are found in the 736. and 1315. positions of the protein sequence.
3. Secretion of PXDN depends entirely on the presence of the secretory signal peptide on the N-terminus of the protein. The intact tertiary structure of its heme binding catalytic domain is necessary and sufficient for the *in vitro* peroxidase activity of PXDN. The C-terminal von Willebrand factor type C domain does not play a role in any of the presently known functions of PXDN. Assembly of the homotrimeric complex is not required for PXDN to localize to the endoplasmic reticulum, but appears to be necessary for both adhesion to the cell surface and optimal collagen IV crosslinking activity of the protein. This suggests that the physiological sites of the enzymatic reaction between collagen IV NC1 domains and PXDN might be the PXDN-rich "hot spots" present on the cell surface, which might also bear an additional, adhesive role.
4. Based on their homology and the conservation of cysteine residues involved in the oligomerization process of PXDN it is likely that PXDNL forms a homotrimeric structure with the same mechanism as PXDN. As seen in the case of its homologue, heterologously expressed recombinant PXDNL displays dual localization and attaches to the cell surface upon secretion from the endoplasmic reticulum. PXDNL's function in the extracellular matrix is hitherto unknown, but since the protein lacks peroxidase activity, it might have a primarily adhesive role.

## LIST OF PUBLICATIONS

**Publications constituting the basis for this dissertation:**

1. **Lázár E**, Péterfi Z, Sirokmány G, Kovács HA, Klement E, Medzihradszky KF, Geiszt M. (2015) Structure-function analysis of peroxidasin provides insight into the mechanism of collagen IV crosslinking. Free Radic Biol Med, 83: 273-82.
   IF: 5,736

2. Péterfi Z, Tóth ZE, Kovács HA, **Lázár E**, Sum A, Donkó A, Sirokmány G, Shah AM, Geiszt M. (2014) Peroxidasin-like protein: a novel peroxidase homologue in the human heart. Cardiovasc Res, 101(3): 393-9.
   IF: 5,940