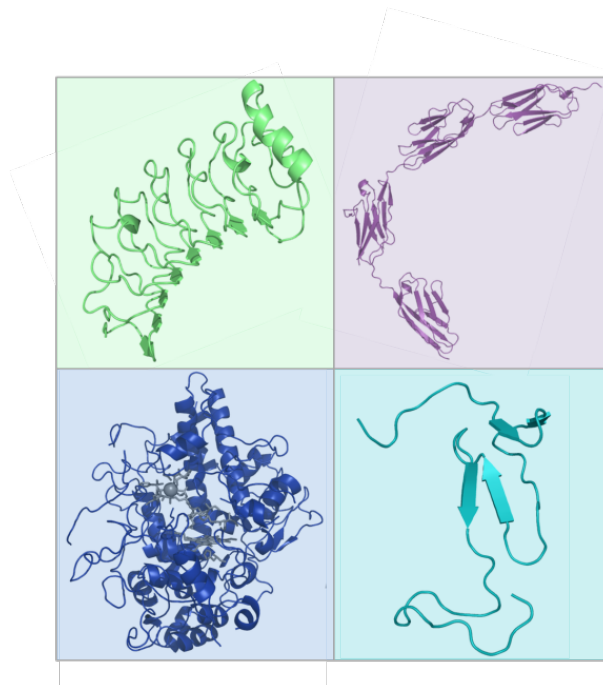


DIPLOMA THESIS

Recombinant expression, biochemical characterization and structural aspects of truncated human peroxidasin 1 variants



Submitted by

Romy-Sophie Katz

Vienna, March 2016

University of Natural Resources and Life Sciences, Vienna

Department of Chemistry, Division of Biochemistry

Protein Biochemistry Group

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to Prof. Dr. Christian Obinger, who gave me the opportunity to do my work in the Protein Biochemistry group and who provided such a pleasant and inspiring atmosphere.

Dr. Martina Paumann-Page was an outstanding supervisor who always answered my questions with endless patience and competence, from the very beginning to the final thesis. Her kindness and her expertise made this a wonderful time. Special thanks also to Dr. Paul Furtmüller who always supported me with his scientific advice.

I also want to thank the whole Bio B working group for their help and for creating a unique atmosphere where it was fun to work in.

Without the never ending love and patience of my parents Angelika and Michael, and their unconditional support in every way I wouldn't have been able to pursue my studies. Also my siblings Angelika and Markus never got tired of encouraging and supporting me. Last but not least I especially want to thank Florian, who motivated me numberless times over the last six years and always stood by my side.

ABSTRACT

Human peroxidasin 1 (hsPxd01) is a multidomain heme containing oxidoreductase, which belongs to the peroxidase-cyclooxygenase superfamily. Its enzymatic peroxidase domain is closely related to that of chordata peroxidases which include the well-studied enzymes myeloperoxidase (MPO), lactoperoxidase (LPO), eosinophil peroxidase (EPO) and thyroid peroxidase (TPO), which comprise Family 1 in this superfamily. The major difference of peroxidasins is the fact that they contain extracellular matrix motifs additionally as they are comprised of an N-terminal leucine-rich repeat domain (LRR), immunoglobulin C-like domains (Ig) and a C-terminal von Willebrand factor C module (VWC) in addition to the catalytic peroxidase domain (POX). As the closely related chordata peroxidases are involved in the innate immune system and the thyroid hormone biosynthesis a similar role was assumed for peroxidasin. The existence of its additional domains further suggested an involvement in the extracellular matrix. Only recently one function of this enzyme was revealed: It takes part in the formation of sulfilimine linkages in the collagen IV network by producing hypohalous acids with bromide representing the substrate of choice for link formation. At the same time, the first evidence of bromide being an essential trace element was provided.

So, human peroxidasin 1 obviously plays an important role in the human body but, however, knowledge about it is still rather poor. As availability of the enzyme is one reason thereof, shorter variants of hsPxd01 were generated in the course of this thesis, omitting distinct domains. Transient expression in mammalian cell culture and subsequent purification steps were established. The constructs could then be characterized using various biochemical methods and the results generated represent a systematic approach towards the characterization of hsPxd01 on an enzymatic level. Furthermore a foundation for future structural investigations was established.

ZUSAMMENFASSUNG

Das humane Enzym Peroxidasin gehört zu den sogenannten Häm-Peroxidasen. Es besteht aus vier verschiedenen Domänen und enthält ein Eisen Ion in seinem aktiven Zentrum. Die Häm-hältige Peroxidasedomäne weist hohe Ähnlichkeit zu anderen humanen Peroxidasen auf, welche bereits eingehend untersucht wurden. Dazu zählen zum Beispiel die Myeloperoxidase oder auch die Laktoperoxidase. Diese Enzyme sind Teil der menschlichen Immunabwehr und die Verwandtschaft zu Peroxidasin legte die Vermutung nahe, dass auch dieses eine ähnliche Rolle im Körper einnehmen könnte. Die drei zusätzlichen Domänen zur Peroxidasedomäne sind allerdings typische Merkmale für Proteine der extrazellulären Matrix. Bis jetzt konnte nur eine Funktion von Peroxidasin mit Sicherheit bestätigt werden, nämlich, dass es die Bildung kovalenter Bindungen zwischen Kollagen IV Untereinheiten katalysiert. Für die Ausbildung dieser Verbindungen wandelt Peroxidasin das Spurenelement Bromid in eine hypobromige Säure um, die dann die Verknüpfung von Aminosäuren ermöglicht. Somit konnte auch der erste Beweis für die essentielle Notwendigkeit von Bromid im Körper geliefert werden.

Da Peroxidasin nach wie vor nicht vollständig untersucht werden konnte und dies zum größten Teil an der mangelnden Verfügbarkeit liegt, wurden im Zuge dieser Diplomarbeit verschiedene verkürzte Konstrukte hergestellt. Das native Protein ist verhältnismäßig groß und dadurch schwierig rekombinant herzustellen. Den Konstrukten fehlen bestimmte Domänen und so konnten sie nicht nur leichter in menschlichen Zellen hergestellt werden, sondern gleichzeitig konnte auch untersucht werden, wie sich ihr biochemisches Verhalten ändert, wenn bestimmte Teile des Ganzen fehlen. Die Konstrukte wurden mit Hilfe von unterschiedlichsten biochemischen Methoden untersucht. Die Resultate ermöglichten eine systematische Charakterisierung und bilden damit eine Grundlage für die weitere Erforschung des Proteins auf enzymatischer wie auch auf struktureller Ebene.

TABLE OF CONTENTS

1	INTRODUCTION	7
1.1	PHYLOGENY OF HEME PEROXIDASES	7
1.1.1	<i>Peroxidase-cyclooxygenase superfamily</i>	<i>7</i>
1.2	PEROXIDASINS	9
1.2.1	<i>Human peroxidasin 1 (hsPxd01)</i>	<i>11</i>
1.2.1.1	Reaction mechanism.....	14
1.2.1.2	Proposed physiological role	16
1.3	RECOMBINANT EXPRESSION OF MAMMALIAN PROTEINS	19
1.3.1	<i>Human embryonic kidney cells (HEK).....</i>	<i>19</i>
2	AIMS OF INVESTIGATION.....	21
3	MATERIAL AND METHODS.....	22
3.1	CLONING OF TRUNCATED HSPXD01 VARIANTS	22
3.1.1	<i>Primer design</i>	<i>27</i>
3.1.2	<i>DNA preparations</i>	<i>28</i>
3.1.2.1	Vector for transfection of HEK 293-6E cells	28
3.1.2.2	Insert DNA amplification by PCR	31
3.1.2.3	Purification of DNA fragments from agarose gel	33
3.1.2.4	Ligation.....	33
3.1.3	<i>Transformation of chemical competent E. coli cells</i>	<i>35</i>
3.1.4	<i>Screening for positive clones</i>	<i>36</i>
3.1.4.1	PCR Screening of transformed E. coli cells.....	36
3.1.4.2	Mini Prep: Plasmid DNA preparation of transformed E. coli cells.....	37
3.1.4.3	Plasmid DNA sequencing.....	39
3.1.5	<i>Preparation of cryo cultures.....</i>	<i>40</i>
3.1.6	<i>Plasmid isolation from E. coli cultures.....</i>	<i>40</i>
3.2	PROTEIN PRODUCTION	41
3.2.1	<i>Human embryonic kidney cells in cell culture</i>	<i>41</i>
3.2.2	<i>Sterile technique</i>	<i>41</i>
3.2.3	<i>Preparation of culture media and additives.....</i>	<i>41</i>
3.2.4	<i>Thawing of HEK 293-6E cells.....</i>	<i>42</i>
3.2.5	<i>Maintenance of HEK cells.....</i>	<i>43</i>
3.2.6	<i>Transient Transfection of HEK 293-6E cells.....</i>	<i>43</i>
3.2.7	<i>Harvest of cells.....</i>	<i>44</i>
3.3	PROTEIN PURIFICATION	45

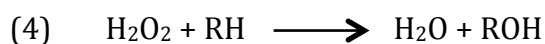
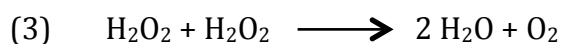
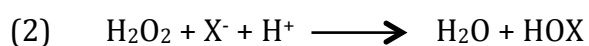
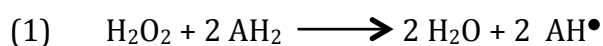
3.3.1	<i>Diafiltration</i>	45
3.3.2	<i>Immobilized metal ion affinity chromatography (IMAC)</i>	45
3.3.3	<i>Concentrating</i>	47
3.3.4	<i>Dialysis</i>	47
3.4	PROTEIN CHARACTERIZATION	48
3.4.1	<i>UV-Vis spectroscopy</i>	48
3.4.2	<i>SDS-PAGE</i>	49
3.4.3	<i>Western Blot</i>	50
3.4.4	<i>Detection of heme by enhanced chemiluminescence</i>	52
3.4.5	<i>Circular dichroism spectroscopy</i>	52
3.4.6	<i>Redox potential</i>	53
3.4.7	<i>Activity measurements</i>	54
3.4.7.1	Peroxidase activity assays	54
3.4.7.2	Halogenation activity assays	55
3.4.8	<i>Polarographical measurement of hydrogen peroxide consumption</i>	57
3.4.9	<i>Pre-steady-state kinetics</i>	58
3.4.9.1	Reaction with hydrogen peroxide – Compound I formation	59
3.4.9.2	Cyanide binding:	60
3.5	STRUCTURAL INVESTIGATIONS	61
3.5.1	<i>Size-exclusion chromatography</i>	61
3.5.2	<i>X-ray crystallography</i>	62
3.5.3	<i>Small-angle X-ray scattering (SAXS)</i>	66
4	RESULTS AND DISCUSSION	67
4.1	CLONING	67
4.1.1	<i>Plasmid preparation</i>	67
4.1.2	<i>Insert preparation</i>	68
4.1.3	<i>Screening</i>	68
4.2	OPTIMIZATION OF PROTEIN PURIFICATION	69
4.2.1	<i>Preparation of supernatant</i>	69
4.2.2	<i>IMAC</i>	70
4.2.3	<i>SDS-PAGE, ECL and Western blot</i>	72
4.3	PROTEIN CHARACTERISATION	73
4.3.1	<i>UV-Vis spectroscopy</i>	73
4.3.2	<i>Comparison of constructs</i>	75
4.3.2.1	SDS-PAGE and ECL	75
4.3.3	<i>Investigations on the C-terminus</i>	77
4.3.4	<i>Circular dichroism spectroscopy</i>	79
4.3.5	<i>Spectroelectrochemistry</i>	81
4.3.6	<i>Activity measurements</i>	83

4.3.6.1	Peroxidase activity	83
4.3.6.2	Halogenation activity	87
4.3.7	<i>Hydrogen peroxide consumption</i>	94
4.3.8	<i>Pre-steady state kinetics</i>	100
4.3.8.1	Compound I formation	100
4.3.8.2	Cyanide binding:.....	103
4.4	STRUCTURAL INVESTIGATIONS.....	106
4.4.1	<i>X-ray crystallography</i>	106
4.4.2	<i>Preliminary SAXS data results</i>	106
5	CONCLUSIONS	108
6	REFERENCES	111

1 Introduction

1.1 *Phylogeny of heme peroxidases*

Peroxidases are oxidoreductases and are able to cleave peroxidic bonds heterolytically. Heme peroxidases contain either heme *b* or posttranslationally modified heme as a cofactor and are among the most common peroxidases (Zamocky *et al.*, 2008). They are found in all kingdoms of life and can catalyse one- or two-electron oxidations mediated by hydrogen peroxide. The turnover of H₂O₂ can lead to different reactions depending on whether one-electron donors (1), two-electron donors (2), catalytic partners (3) or oxygen functionalities (4) are involved in the reaction:



Numerous investigations and classifications today suggest the presence of four distinct peroxidase superfamilies which arose independently: peroxidase-catalase, peroxidase cyclooxygenase, peroxidase-chlorite dismutase and peroxidase-peroxygenase superfamily (Zámocký *et al.*, 2015).

1.1.1 Peroxidase-cyclooxygenase superfamily

Peroxidasins belong to the peroxidase-cyclooxygenase superfamily. Characteristic for members of this superfamily is their involvement in the innate immune system, which includes peroxidases from mammals as well as from bacteria (Zámocký and Obinger, 2010).

The superfamily is divided into 7 families (see Figure 1-1), whereby numerous members of them are multidomain proteins with a predominantly alpha helical fold of a heme containing core.

Peroxidasins comprise Family 2 and are closely related to Family 1, the chordata peroxidases and Family 3, the peroxinectins.

The chordata peroxidases (Family 1) include myeloperoxidase (MPO), eosinophil peroxidase (EPO), lactoperoxidase (LPO) and thyroid peroxidase (TPO). The comparison of these enzymes to peroxidasins is helpful as they are rather well studied. MPO, EPO and LPO are part of the innate immune system as they produce hypohalous acids to kill pathogens while TPO is involved in the synthesis of thyroid hormones.

In Figure 1-1 it can be clearly seen that peroxidasins and chordata peroxidases evolved from a common ancestor (Zamocky *et al.*, 2008).

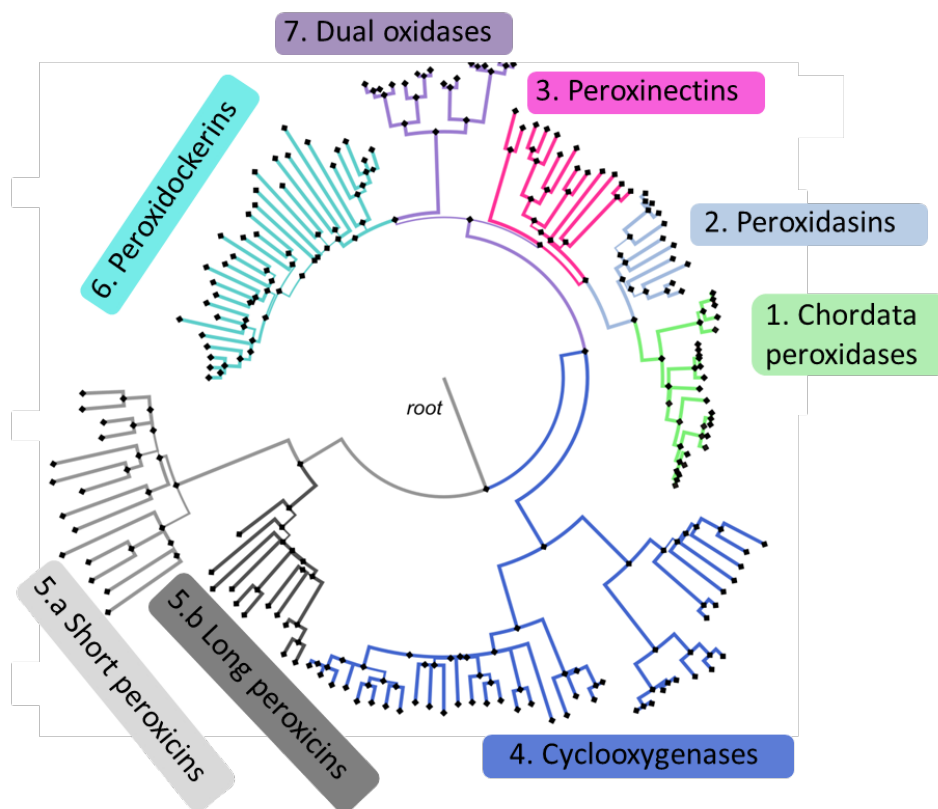


Figure 1-1: Overview of phylogeny of peroxidase-cyclooxygenase superfamily (Zamocky *et al.*, 2008)

1.2 Peroxidasins

As mentioned above, peroxidasins (EC 1.11.1.7) comprise Family 2 of the peroxidase-cyclooxygenase superfamily. These multidomain enzymes contain a heme in their catalytically active peroxidase domain. This peroxidase domain (POX) further shows high homology to the human chordata enzyme lactoperoxidase (LPO) in terms of amino acid sequence and active site architecture (Soudi *et al.*, 2012).

Peroxidasins were first described for *Drosophila* in 1994 by Nelson and colleagues (Nelson *et al.*, 1994). A closer look at the phylogeny of peroxidasins is shown in Figure 1-2 whereby the division into 5 subfamilies is visible.

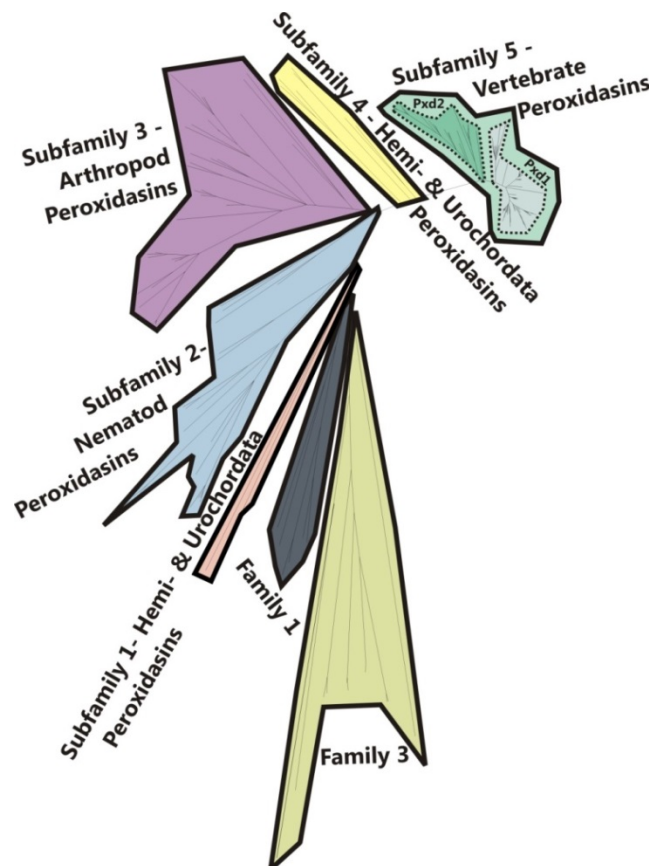


Figure 1-2: Subfamilies of Family 2 (peroxidasins)

As seen in Figure 1-2, peroxidasins are found in invertebrates – which are represented in subfamilies 1-4 – as well as in vertebrates, which comprise subfamily 5.

To further characterize vertebrate peroxidasins (subfamily 5) Figure 1-3 shows that they can be divided into two different branches designated as peroxidasin 1 and peroxidasin 2. The existence of two different human variants was confirmed about five years after the first description of peroxidasin (Horikoshi *et al.*, 1999).

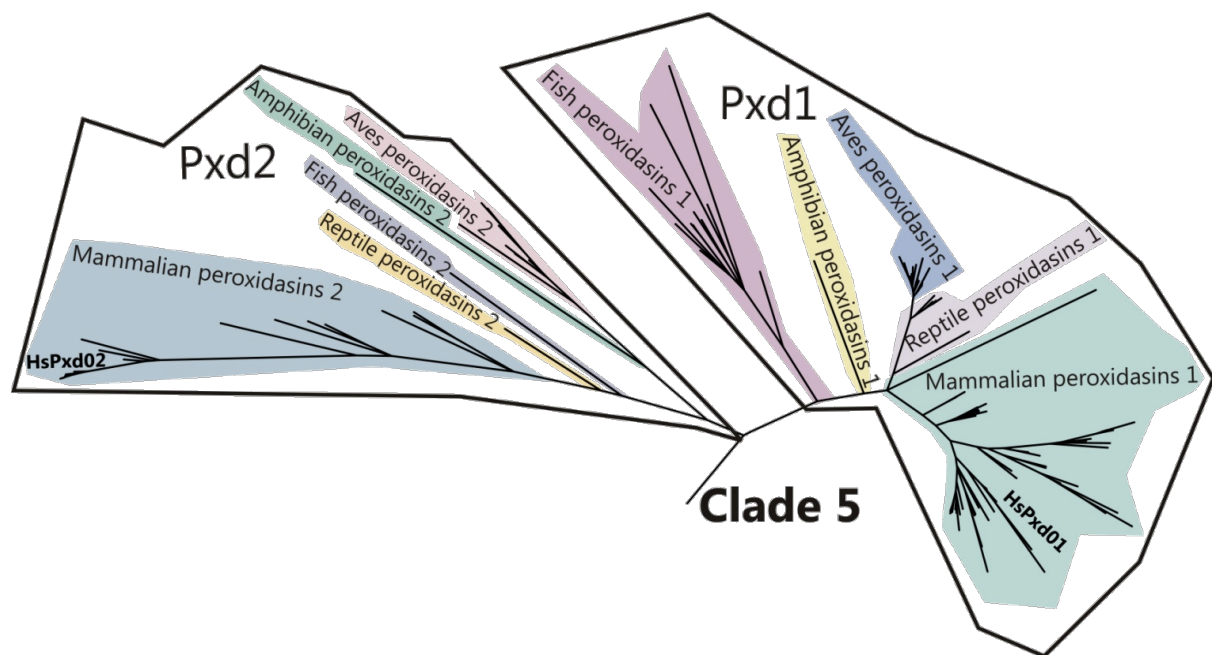


Figure 1-3: Separation of subfamily 5 into two branches: peroxidasin 1 (Pxd1) and peroxidasin 2 (Pxd2).

Those two forms of human peroxidasin share a sequence identity of 58 % and a high sequence similarity of 72 % (Péterfi *et al.*, 2014). Human peroxidasin 1 (hsPxd01) is located on chromosome 2p25; hsPxd02 is located on chromosome 8q11. They both consist of the same four domains (described in chapter 1.2.1) as predicted by sequence analysis but have striking differences in the active site of their peroxidase domains. This leads to the conclusion that they might have different biological roles (Soudi *et al.*, 2012).

1.2.1 Human peroxidasin 1 (hsPxd01)

At the moment, human peroxidasin 1 is studied more extensively than hsPxd02. As mentioned before both comprise of 4 distinct domains. Figure 1-4 shows the composition of hsPxd01. In this figure, also the signal peptide is included. It facilitates the secretion of the protein to the extracellular space but it is not present in the mature form or in the recombinant protein.

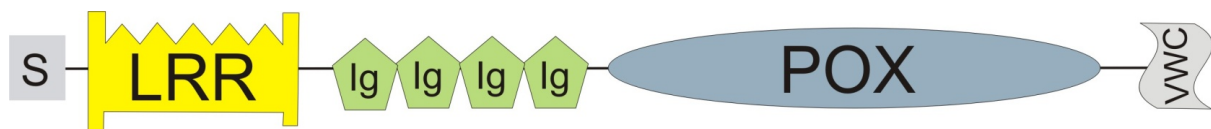


Figure 1-4: Composition of human peroxidasin: from left to right: signal peptide (S) - leucine rich repeat domain (LRR), Ig domains (Ig), peroxidase domain (POX), von Willebrand factor type C domain (VWC).

The N-terminal leucine-rich repeat domain (LRR) consists of five typical repeats which are flanked by N- and C-terminal capping regions. Following this, four immunoglobulin C like motifs comprise the Ig domain (Ig). These immunoglobulins consist of seven beta-strands each. Next the catalytically active peroxidase domain (POX) is located. This domain possesses a mainly alpha helical fold and contains a modified heme at its active site. Finally, at the C-terminus the von Willebrand factor type C domain is located. This domain is also sometimes referred to as chordin-like cysteine-rich (CR) repeats. The three non-catalytic domains, LRR, Ig and VWC, are extracellular matrix motifs and are most likely necessary for protein-protein interactions (Zamocky *et al.*, 2008) (Soudi *et al.*, 2012). Human peroxidasin 1 further is a highly glycosylated enzyme. There are 11 putative N-glycosylation sites throughout the protein, whereby seven of them are predicted to be on the peroxidase domain (Soudi *et al.*, 2015).

As there don't exist any crystallographically data at the moment the different domains could only be modelled compared to similar enzymes like human LPO or goat LPO (gLPO) and MPO. These models are shown in Figure 1-5 (Soudi *et al.*, 2012).

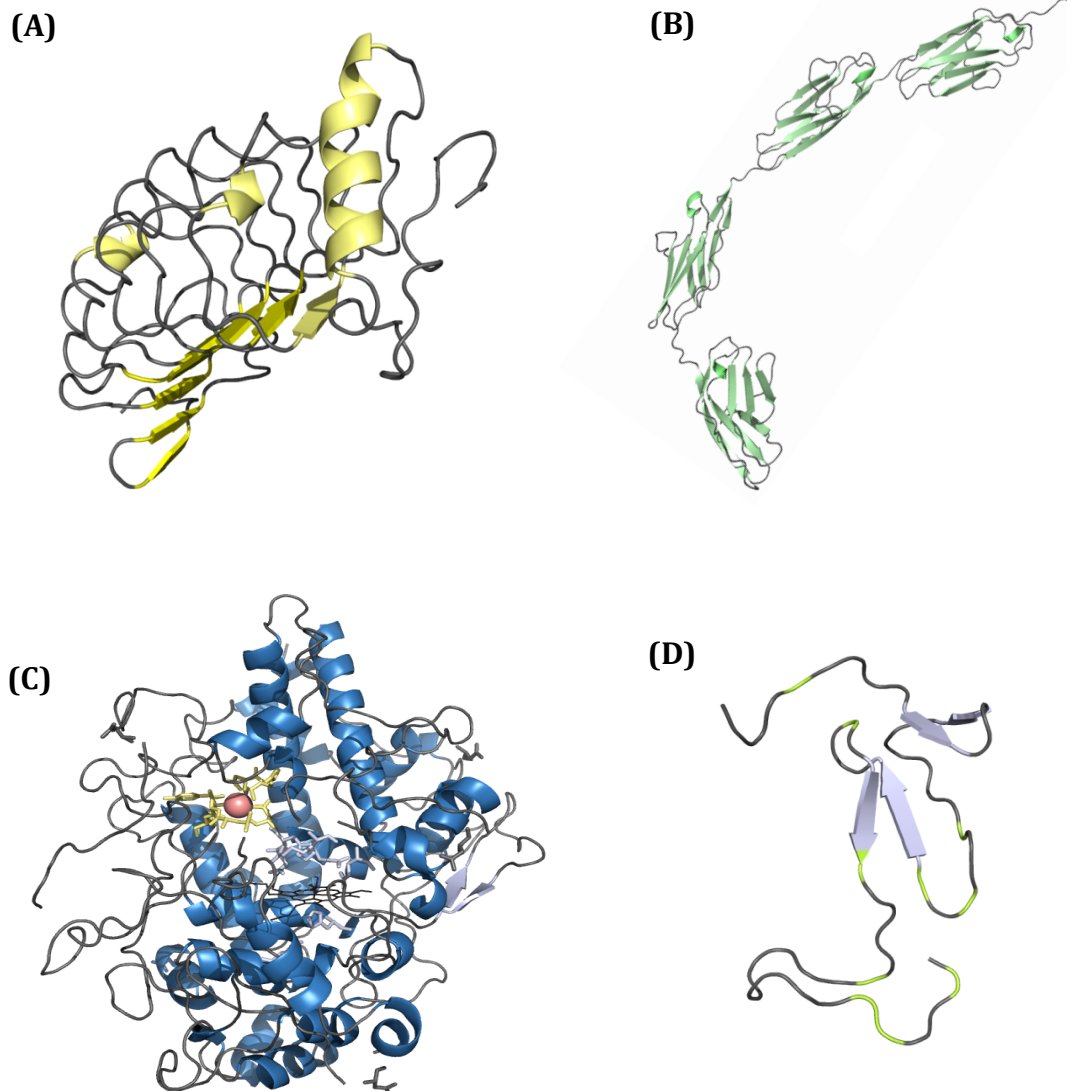


Figure 1-5: Models of the four distinct domains of human peroxidase 1: (A) leucine-rich repeat domain, (B) Ig domains, (C) peroxidase domain containing the heme, (D) von Willebrand factor C domain.

The prosthetic heme group

As mentioned, the heme is located in the peroxidase domain and is of crucial importance for the enzymatic activity. This prosthetic group of heme peroxidases consists of a coordinated iron atom in the centre of an organic ring structure called

porphyrin. Therefore it is described to be a ferri-protoporphyrin IX derivative. Cheng and colleagues described it to be covalently bound to the protein. These links are formed autocatalytically (Cheng *et al.*, 2008). Ester linkages are established between the 1- and 5-methyl carbons and a glutamate residue (Glu980) and an aspartate residue (Asp826). This binding pattern is conserved in all members of the family. The iron atom in the centre is coordinated to four pyrrole nitrogen atoms of the heterocyclic organic protoporphyrin ring as well as to a proximal histidine (Zederbauer *et al.*, 2007).

Figure 1-6 shows a model of the active site structure of the peroxidase domain of human peroxidasin (blue) based on the crystal structure of goat LPO (red). The heme is depicted in grey.

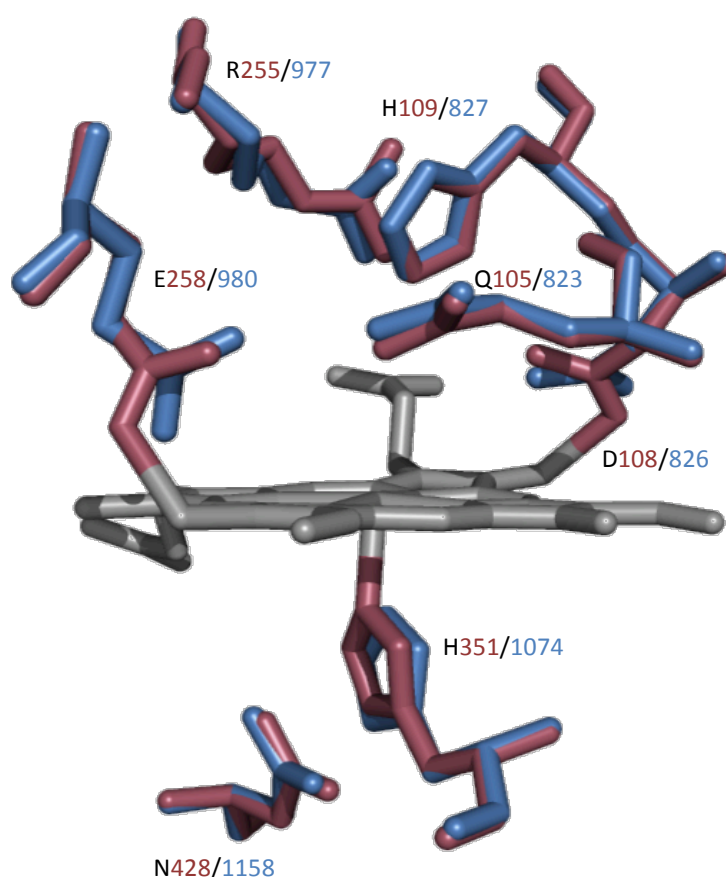


Figure 1-6: Active site structure of the peroxidase domain of hsPxd01 modelled on the crystal structure of goat LPO (PDB code 2R5L): numbering of the active site residues corresponds to mature LPO (red), whereas numbering in blue corresponds to full length hsPxd01 (PyMol).

The sequence alignment of the human peroxidasin 1 peroxidase domain (Con 5N) and gLPO is depicted below. Amino acid residues of the highly conserved active site are highlighted in red, other highly conserved and similar residues are highlighted in blue and yellow.

Highly conserved residues Residues with similarity Highly conserved active site residues

Studies performed so far showed that human peroxidasin 1 shows peroxidase activity and halogenation activity. It can catalyse one-electron as well as two-electron oxidation reactions (Soudi *et al.*, 2012). Oxidation of halides is essential for the physiological role of peroxidasin as described in chapter 1.2.1.2.

and one water molecule is released. The second oxygen atom gets coordinated to the heme and Compound I [oxoiron(IV) porphyrin radical] is formed. The formation of Compound I (Reaction 1 in Figure 1-7) is the first step for both the peroxidase cycle and the halogenation cycle (Furtmüller *et al.*, 2006) (Jantschko *et al.*, 2004).

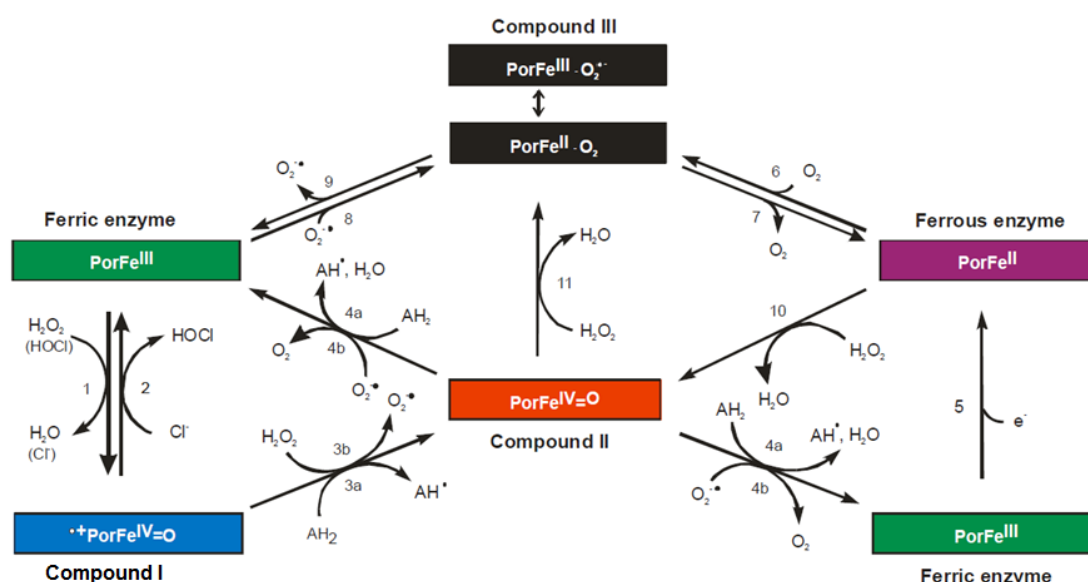
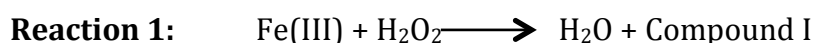


Figure 1-7: General peroxidase reaction scheme: redox pathways (Jantschko *et al.*, 2004)



This formation of Compound I is a very fast reaction and takes place within a few milliseconds (Zámocký and Obinger, 2010).



In the peroxidase cycle Compound I is reduced to Compound II [oxoiron(IV)] which is shown in Reaction 3a in Figure 1-7, which is further reduced back to the native ferric form in another one-electron reduction reaction (Reaction 4a in Figure 1-7) (Jantschko *et al.*, 2004).

In the case of peroxidase, the halogenation cycle is of special interest as the oxidation of halides is seen as a major part of its physiological function (see chapter 1.2.1.2). Here Compound I mediates the two-electron oxidation of halides and pseudo-halides

whereby it is reduced back to its ferric form (Reaction 2 in Figure 1-7) (Jantschko *et al.*, 2004).

Another important property of peroxidasin and heme peroxidases in general is the binding of ligands. If for example cyanide binds to the ferric high-spin heme iron of peroxidasin, a low-spin complex is formed. This binding reaction is used as an indicator for substrate accessibility to the active site of peroxidases (Furtmüller *et al.*, 2006).

1.2.1.2 Proposed physiological role

Peroxidasin differs from the related chordata enzymes as it contains extracellular matrix motifs in addition to its peroxidase domain. That's why it is thought that it could have a special role in the formation of extracellular matrix (Lázár *et al.*, 2015). Studies like those from Cheng and colleagues or Péterfi and colleagues proofed that mRNA encoding human peroxidasin 1 occurs in many different tissues like heart, colon, spleen, kidney, skeletal muscle, liver and many more. In contrast, mRNA encoding hsPxd02 was only found in the adult heart so far (Cheng *et al.*, 2008) (Péterfi *et al.*, 2014). This once more confirms the assumption that the physiological roles of hsPxd01 and hsPxd02 are different ones.

Secretion and incorporation of human peroxidasin 1 into the extracellular matrix was already shown in 2009 but the underlying reason could not be identified back then (Péterfi *et al.*, 2009). An important step for the elucidation of the function was accomplished by Bhavé and colleagues in 2012 when they could show that hsPxd01 is directly involved in the formation of a sulfilimine link in the collagen IV network. Collagen IV is the main constituent of basement membranes and provides structural integrity to tissues. Peroxidasin is responsible for catalysis of this unique sulfilimine link in the non-collagenous (NC1) domain of collagen IV. Figure 1-8 shows the crystal structure of two opposing NC1 protomers (PDB code 1LI1). They are situated at the C-termini of collagen IV monomers. Mammalian collagen IV is derived from six α -chain polypeptides, whereby the NC1 domain of each chain is about 230 amino acids long (Abreu-Velez and Howard, 2012). Noteworthy is that 2 specific amino acids are involved in the formation of this sulfilimine linkage, namely methionine 93 and lysine 211.

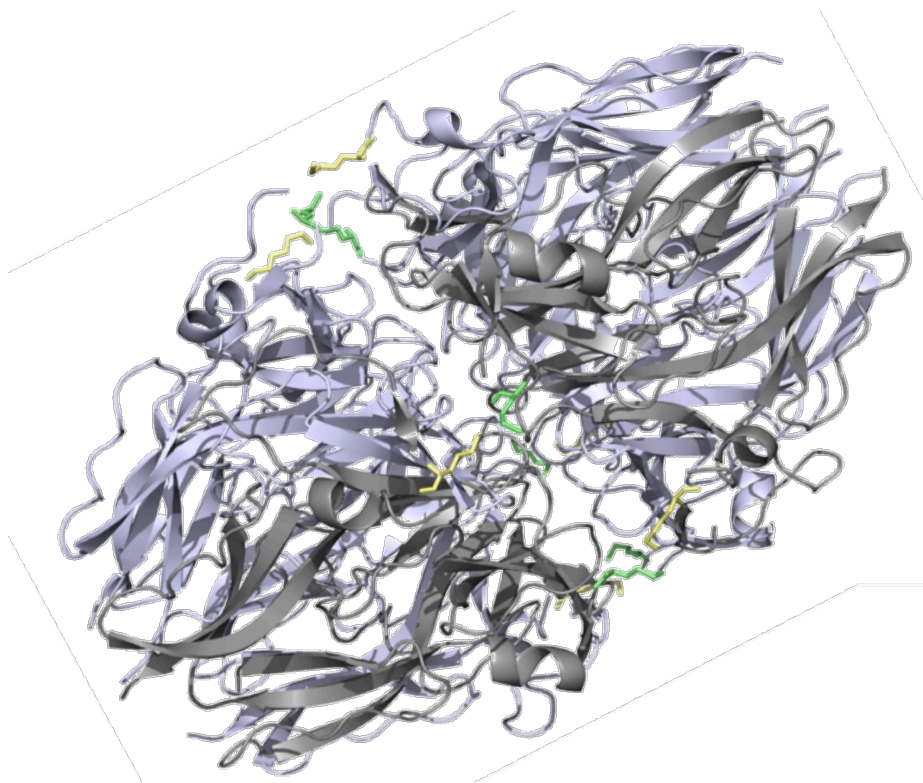


Figure 1-8: Two of the NC1 trimers associate with each other, resulting in a hexameric structure: green: Met93 residues; yellow: Lys211 residues

These two amino acids are oxidized by hypohalous acids generated by human peroxidase 1. Therefore McCall and colleagues could verify the fact that bromide is necessary for this reaction. In the presence of peroxidase bromide is oxidized to hypobromous acid which is further needed for the synthesis of the linkage. Thus, not only a biosynthetic role of human peroxidase 1 in the extracellular matrix formation was demonstrated but also the first evidence for bromide being an essential trace element was revealed. Peroxidase might also be responsible for the formation of hypochlorous acid, however, to a much smaller extent. Figure 1-9 shows the linkage formation in more detail. When the respective hypohalous acid reacts with the methionine sulfur atom, a halosulfonium cation intermediate is formed, which subsequently reacts with the amine nucleophile of the lysine residue to form the sulfilimine link (Bhave *et al.*, 2012) (McCall *et al.*, 2014).

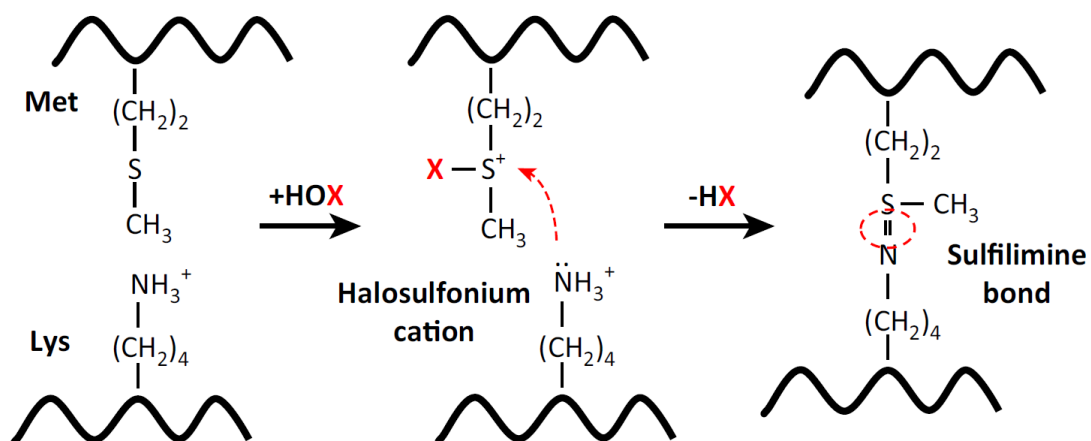


Figure 1-9: Sulfilimine link formation between the amino acid residues Met93 and Lys211

As hypochlorous acid alone would not be sufficient for this link formation due to mainly sulfoxide generation, a deficiency of bromide leads to basement membrane defects. There is a couple of well characterized pathologies which are caused by basement membrane alterations, one of them is the Goodpasture's disease. In this autoimmune disease, antibodies which recognize and bind to parts of the NC1 domain of collagen IV, lead to inflammation and even glomerulonephritis which means the disruption of the glomerular filtration barrier in the kidneys. This is usually prohibited by the structural constraint represented by sulfilimine crosslinked NC1 domains (Hellmark and Segelmark, 2014).

1.3 Recombinant expression of mammalian proteins

Recombinant expression of proteins means the production of a protein in an organism which originally did not carry the distinct gene encoding for this protein. This method makes it possible to produce large amounts of proteins and is especially important for those proteins which cannot be gained out of natural resources or if this process is combined with excessive effort or costs.

As posttranslational modifications of mammalian proteins are often much more complex than bacteria are able to perform, methods for transient transfection of animal cell lines are of crucial importance today. Especially human embryonic kidney cells as well as Chinese hamster ovary cells are in common use for large scale production today (Baldi *et al.*, 2005).

The procedure of transient gene expression makes rapid production in relatively large amounts possible and is therefore a very important method in animal cell culture. In contrast to production using a stable cell line, which can take months to be established, time from transfection to harvest is only a few days (Wurm, 2004).

Despite all these advantages one still has to consider that recombinant protein production in mammalian systems is very expensive compared to the expression in yeast or bacteria. Besides, cultivation is also labour-intensive and often it is quite difficult to establish a cell line in serum-free media.

1.3.1 Human embryonic kidney cells (HEK)

The HEK cell line was established in the 1970s and is used as a common expression tool ever since. Although it is primarily appreciated as a reliable system for transient transfections, also stable clones were produced successfully. For highly glycosylated proteins like peroxidase, the feature of this cell line to carry out almost all post-translational modifications is especially valuable.

When talking about “HEK cells” we actually mean the most common type, the so-called HEK293 cell line. This evolved when human embryonic kidney cells were transformed with parts of DNA of the human adenovirus type 5 (Ad5) which were introduced into

the chromosome of HEK cells. These cells express the E1A adenovirus gene which further takes part in activation of different viral promoters. This mechanism is responsible for the production of high levels of protein.

By further transformation, the HEK293-6E cell line was created. This type additionally expresses the Epstein-Barr virus nuclear antigen 1 (EBNA1). Thus, episomal replication of plasmids carrying the *oriP* origin of replication is possible. Further, transcription of transgenes can be enhanced when carrying the human cytomegalovirus (CMV) promoter. Another important feature of HEK cells is that they can grow in suspension and without serum and that protein expression is comparatively fast (Thomas and Smart, 2005).

2 Aims of investigation

The heme containing enzyme human peroxidase 1 is characterized only partly by now. One reason therefore is that so far it's not possible to purify the protein from natural resources and that the recombinant production of the full length protein has been proven difficult as well. Only very low amounts could be expressed and purified. To learn more about this quite large enzyme, shorter truncated variants should be expressed and reveal among others the functions of the domains.

So the first goal was to establish a high-yield production and purification protocol or to be precise to optimize the already existing procedure. As a start the already established constructs were investigated.

Further, some more different constructs were cloned using new cloning strategies. As many reasonable variants as possible were produced and the expression was further tested at small scale.

Another important goal was to characterize the biochemical properties of the variants using different methods to generate a broad spectrum of data. Activity measurements were performed in depth and compared with the rather well characterized chordata enzymes like LPO and MPO.

Furthermore different experiments were performed to elucidate structural aspects of the enzyme as this is always an important part of a comprehensive protein characterization. Some preliminary data was created which can be helpful for further investigations.

3 Material and methods

3.1 Cloning of truncated hsPxd01 variants

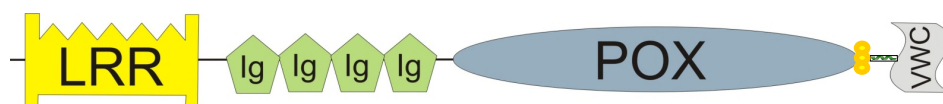
Human peroxidasin 1 (hsPxd01) is a multidomain enzyme containing a heme in its active site. The recombinant full length enzyme, containing all of these domains except for the endogenous signal peptide, could only be expressed at relatively low yields so far. That's why it was decided to work with truncated variants, omitting distinct domains of the protein. These shortened variants can be used as model enzymes for full length hsPxd01. Furthermore, it is possible to study the impact of the respective domains on enzymatic activity, covalent heme attachment and structural organization. Some of these constructs were already established before and some were cloned and produced in the course of this thesis. They were named systematically Con 1N to Con 5N whereas "N" stands for the N-terminal hexa-His-tag. Most of the truncated variants could be expressed at considerably higher rate than full length hsPxd01.

The C-terminus of hsPxd01:

It was already known that human peroxidasin 1 forms a redox sensitive homotrimer. To elucidate the exact sites of oligomerization additional variants were created varying in the length of their C-terminus. The linker region – connecting the peroxidase domain with the C-terminal VWC domain - contains a short stretch with three cysteine residues (depicted in the figures below as yellow circles) followed by an alpha helical region (depicted as green helix). According to the amino acid sequence one variant was designed to only contain the three cysteine residues (cys) and the other one included the amphipathic helical sequence as well (helix).

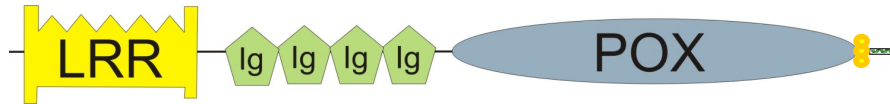
- **Construct 1N (Con 1N)**

Construct 1N is identical with the full length enzyme as it consists of the leucine-rich repeat domain, the four Ig domains, the catalytically active peroxidase domain as well as the von Willebrand factor type C domain. It has a theoretical monomeric molar mass of 164 kDa. Detailed sequence information of the constructs can be read off in Table 3-2.



- **Construct 1N_helix (Con 1N_helix)**

This construct is lacking the VWC domain but still contains the linker region including the helical sequence. A theoretical monomeric molar mass was determined to be 156 kDa.



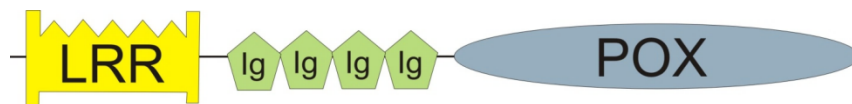
- **Construct 2N (Con 2N)**

In Construct 2N the leucine-rich repeat domain was omitted, resulting in a monomeric molar mass of 138 kDa.



- **Construct 3N (Con 3N)**

Construct 3N is lacking the C-terminal von Willebrand factor type C domain as well as the linker region and therefore ends just after the peroxidase domain (145 kDa).



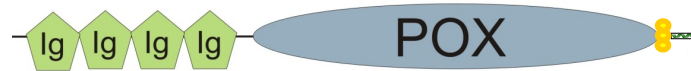
- **Construct 4N (Con 4N)**

This construct consists of the four Ig domains and the peroxidase domain and has a monomeric molecular mass of 121 kDa. It has already been established and an *E. coli* cryo culture was available for plasmid DNA preparation.



- **Construct 4N_helix (Con 4N_helix)**

Construct 4N helix contains the 4 Ig domains and the peroxidase domain like Construct 4N. Additionally, like Construct 1N_helix and Construct 5N_helix, this variant includes the amphipathic helical sequence at the C-terminus. Its monomeric molar mass amounts to 131 kDa.



- **Construct 5N (Con 5N)**

Con 5N comprises only the peroxidase domain and like Con 4N this variant was already established and available (81 kDa).



- **Construct 5N_VWC (Con 5N_VWC)**

Construct 5N_VWC was the initial variation of Con 5N with the C-terminal addition of the linker region after the peroxidase domain and the von Willebrand factor type C. As a monomer this variant is 98 kDa.



- **Construct 5N_cys (Con 5N_cys)**

In Con 5N_cys, which is another variation of Con 5N, the peroxidase domain was elongated by the short part of the linker region which contains the three cysteine residues, which results in a theoretical monomeric molar mass of 81 kDa.



- **Construct 5N_helix (Con 5N_helix)**

Construct 5N_helix is an extended version of Con 5N_cys. It additionally contains the amphipathic helical sequence and ends before the VWC domain. Its monomeric size accounts for 91 kDa.



All constructs that possess the linker region after the peroxidase domain form trimers.

Table 3-1: Information about hsPxd01 regions

hsPxd01 structure	Amino acid sequence	Amino acids	Cysteines	Glycosylation sites
Signal sequence	1 – 26	26	2	1
LRR region	27 – 245	219	8	-
IgG regions	246 – 619	374	8	1
POX domain	620 – 1314	695	15	7
Linker region	1315 - 1412	98	3	1
VWC domain	1413 – 1471	59	10	1
total	1479		46	11

Table 3-2: Sequence information of the different constructs

hsPxd01 variant	Amino acids	From – to		Cysteines	Glycosylation sites
Con 1N	1445	Val27	Leu1479	44	10
Con 1N_helix	1385	Val27	Ser1411	34	9
Con 2N	1226	Pro246	Leu1479	36	10
Con 3N	1288	Val27	Asp1314	31	8
Con 4N	1069	Pro246	Asp1314	23	8
Con 4N_helix	1166	Pro246	Ser1411	26	9
Con 5N	695	Gly620	Asp1314	15	7
Con 5N_VWC	852	Gly620	Leu1471	28	9
Con 5N_cys	708	Gly620	Ala1327	18	7
Con 5N_helix	792	Gly620	Ser1411	18	8

Amino acid sequence of hsPxd01:

```

10      20      30      40      50      60      70      80
MAKRSRGPGR RCLLALVLFC AWGTLAVVVAQ KPGAGCPSRC LCFRTTVRCM HLLLEAVPAV APQTSILDLR FNRIREIQPG

90      100     110     120     130     140     150     160
AFRRRLNLNT LLLNNNQIKR IPGAFEDLE NLKYLYLYKN EIQSIDRQAF KGLASLEQLY LHFNQIETLD PDSFQHLPKL

170     180     190     200     210     220     230     240
ERLFLHNNRI THLVPQTFNH LESMKRLRLD SNTLHCDCEI LWLADLLKTY AESGNAQAAA ICEYPRRIQG RSVATITPEE

250     260     270     280     290     300     310     320
LNCERPRITS EPQADAVTSG NTVYFTCRAE GNPKPEIWL RNNNELSMKT DSRLNLLDDG TLMIQNTQET DQGIYQCMAK

330     340     350     360     370     380     390     400
NVAGEVKTQE VTLRYFGSPA RTFTVIQPQN TEVLVGESVT LECSATGHPP PRISWTRGDR TPLPVDPRVN ITPSGGLYIQ

410     420     430     440     450     460     470     480
NVVQGDSEYE ACSATNNIDS VHATAFIIVQ ALPQFTVTPQ DRVVIEGQTV DFQCEAKGNP PPVIAWTKGG SQLSVDRRHL

490     500     510     520     530     540     550     560
VLSSGTLRIS GVALHDQGYE CQAVNIIGS QKVVAHLTVQ PRVTPVFASI PSDTTVEVGA NVQLPCSSQG EPEPAITWNK

570     580     590     600     610     620     630     640
DGVQVTESGK FHISPEGFLT INDVGPADAG RYECVARNTI GSASVSMVLS VNVDPVSRNG DPFVATSIVE AIATVDRAINN

650     660     670     680     690     700     710     720
STRTHLFDSR PRSPNDLLAL FRYPRDPYTV EQARAGEIFE RTLQLIQEHV QHGLMVDLNG TSYHYNDLVS PQYLNLIANNL

730     740     750     760     770     780     790     800
SGCTAHRRVNN CSDMCFFHQK YRTHDGTCNN LQHPMWGASL TAFERLLKSV YENGFNTPRG INPHRLYNGH ALPMPRLVST

810     820     830     840     850     860     870     880
TLIGTETVTP DEQFTHMLMQ WGQFLDHDLD STVVALSQAR FSDGQHCSNV CSNDPPCFSV MIPPNDSRAR SGARCMFFVR

890     900     910     920     930     940     950     960
SSPVCSGGMT SLLMNSVYPR EQINQLTSYI DASNVYGSTE HEARSIRDLA SHRGLLRQGI VQRSGKPLLP FATGPPTECM

970     980     990     1000    1010    1020    1030    1040
RDENESPIPC FLAGDHRANE QLGLTSMHTL WFREHNRIAT ELLKLNPHWD GDTIYYETRK IVGAEIQHIT YQHWLPKILG

1050    1060    1070    1080    1090    1100    1110    1120
EVGMRTLGEY HGYDPGINAG IFNAFATAAF RFGHTLVNPL LYRLDENFQP IAQDHLPLHK AFFSPFRIVN EGGIDPLLRG

1130    1140    1150    1160    1170    1180    1190    1200
LFGVAGKMRV PSQLNTELT ERLFSMAHTV ALDLAAINIQ RGRDHGIPPY HDYRVYCNLS AAHTFEDLKN EIKNPEIREK

1210    1220    1230    1240    1250    1260    1270    1280
LKRLYGSTLN IDLFPALVE DLVPGSRLGP TLMCLLSTQF KRLRDGDRLW YENPGVFSPA QLTQIKQTSI ARILCDNADN

1290    1300    1310    1320    1330    1340    1350    1360
ITRVQSDVFR VAEFPHGYGS CDEIPRVDLR VWQDCCEDCCR TRGQFNAFSY HFRGRRSLEF SYQEDKPTKK TRPRKIPSVG

1370    1380    1390    1400    1410    1420    1430    1440
RQGEHLSNST SAFSTRSDAS GTNDFREFVL EMQKTITDLR TQIKKLESRL STTECVDAGG ESHANNTKWK KDACTICECK

1450    1460    1470
DGQVTCFVEA CPPATCAVPV NIPGACCPVCLQKRAEEKP

```

Signal peptide LRR domain IgG domains propeptide homologue of POX
 POX domain "cys" extension "helix" extension VWC domain

^N predicted glycosylation sites

3.1.1 Primer design

The primers were designed for cloning with *Gibson Assembly Cloning Kit* from New England Biolabs which ligates a distinct insert with a linearized vector. The advantage of this method is that no restriction sites are necessary. Insert and vector need overlapping homologous ends (25-80 bp), which were created on the insert with the primers (see Figure 3-3 *Gibson Cloning*). The forward primers further contained the N-terminal hexa-histidine tag. Forward and reverse primers comprised approximately 25 bases which primed on the DNA sequence of interest.

The sequence of each primer was analysed using the program *Gene Runner*.

All primers were ordered at Sigma-Aldrich (scale: 0.025; purified by HPLC) and were resuspended in the appropriate amount of dH₂O as listed on the technical datasheet to obtain a 100 µM solution.

Since construct 4N and construct 5N were already successfully cloned only the following primers (Table 3-3) had to be designed.

Table 3-3: Primer: all primers were designed for cloning with Gibson Assembly and the pTT5+ vector; grey: plasmid; red: hexa-histidine tag; blue: stop codon; black: hsPxd01 construct sequence; dark red: silent mutations to improve primer properties.

Name		Primer sequence 5' → 3'
1	ptt5+_1N_3N_his6 fw_pr_gib	GCTCTGGGTTCCAGGTTCCACTGGCCATCATCACCATCACCATGTGGTGGCCAGAAGCCGGGCGCAGGG
2	ptt5+_Con4N_his6 fw_pr_gib	GCTCTGGGTTCCAGGTTCCACTGGCCATCATCACCATCACCATCCACGAATCACCTCCGAGCCCC
3	ptt5+_5N_his+VWC fw pr gib	GCTCTGGGTTCCAGGTTCCACTGGCCATCATCACCATCACCATGGAGATCCGTTTGTAGCTACCTCCATCG
4	1N-His_rev_pr_gibson	GCCAGAGGTCGAGGTCGGGGGATCCTTATCAGGGCTTTTCCTCCGCCCTCTTCTGTAAGCAGACTGG
5	3N-His_rev_pr_gibson	GCCAGAGGTCGAGGTCGGGGGATCCTTATCAGTCCTGCCACACCCGGAGGTCCACCCTGGGG
6	5N-His_helix_rev_pr_gibson	GCCAGAGGTCGAGGTCGGGGGATCCTTATCAACTGAGCCGTGATTCAAGTTTCTTTATC
7	5N-His_CYS_rev_pr_gibson	GCCAGAGGTCGAGGTCGGGGGATCCTTATCAGGCATTGAACTGCCCCCTGGTCTACAGTCTTCACAGC

With these primers the following new variants were cloned (Table 3-4).

Table 3-4: Primers used for distinct constructs: different variants were cloned using the distinct primer pairs.

Construct	Forward primer		Reverse primer	
Con 1N	1	ptt5+_1N_3N_his6 fw_pr_gib	4	1N-His_rev_pr_gibson
Con 1N_helix	1	ptt5+_1N_3N_his6 fw_pr_gib	6	5N-His_helix_rev_pr_gibson
Con 2N	2	ptt5+_Con4N_his6 fw_pr_gib	4	1N-His_rev_pr_gibson
Con 3N	1	ptt5+_1N_3N_his6 fw_pr_gib	5	3N-His_rev_pr_gibson
Con 4N	Already established			
Con 4N_helix	2	ptt5+_Con4N_his6 fw_pr_gib	6	5N-His_helix_rev_pr_gibson
Con 5N	Already established			
Con 5N_cys	3	ptt5+_5N_his6+VWC fw pr gib	7	5N-His_CYS_rev_pr_gibson
Con 5N_helix	3	ptt5+_5N_his6+VWC fw pr gib	6	5N-His_helix_rev_pr_gibson
Con 5N_VWC	3	ptt5+_5N_his6+VWC fw pr gib	4	1N-His_rev_pr_gibson

3.1.2 DNA preparations

3.1.2.1 Vector for transfection of HEK 293-6E cells

For the transient transfection of all constructs described here, a modified pTT5 vector, namely pTT5+, was used. In the modified plasmid the multiple cloning site of the pTT5 vector was removed using the *EcoRI* and *BamHI* restriction sites and replaced by a synthesized DNA fragment for ease of cloning of previously cloned constructs (see Figure 3-1).

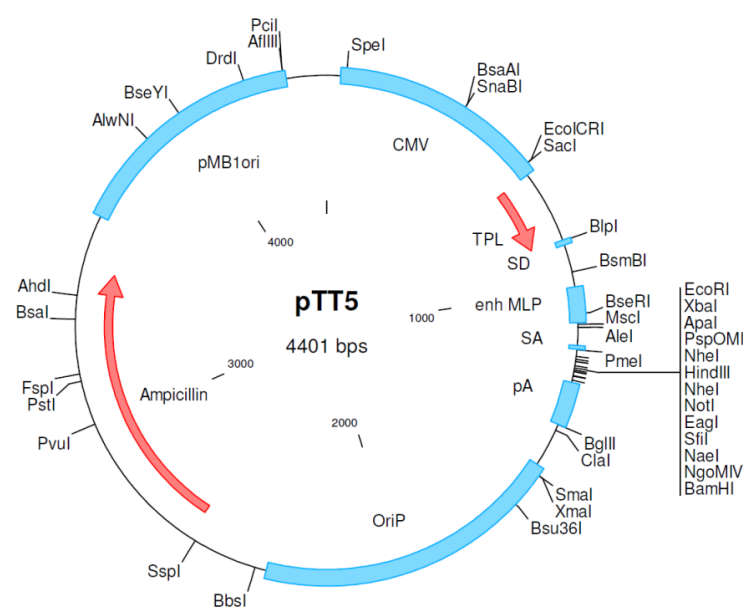


Figure 3-1: pTT5 vector for transient transfection (NRC biotechnology research institute)

As seen in Figure 3-1, this vector possesses different elements:

- a multiple cloning site (MCS) containing recognition sequences for many different restriction enzymes
- a constitutive promotor of human cytomegalovirus (CMV) which is controlling the MCS
- the Epstein-Barr virus origin of replication (*oriP*)
- an ampicillin resistance
- pMB1 origin of replication, necessary for amplification of the vector in bacteria

As HEK293-6E cells stably express the Epstein-Barr virus nuclear antigen 1 (EBNA1) they have the ability for episomal replication of the *oriP* containing pTT5 vector and enhanced transcription of CMV promotor driven transgenes.

Vector preparation for Gibson cloning:

The pTT5+ vector containing the DNA sequence of previously cloned construct 5N (Con 5N, Figure 3-2) was used for preparing the linearized vector for Gibson cloning of the new constructs. The Con 5N insert of this vector was removed by digestion with the restriction enzymes *NheI* and *NotI*.

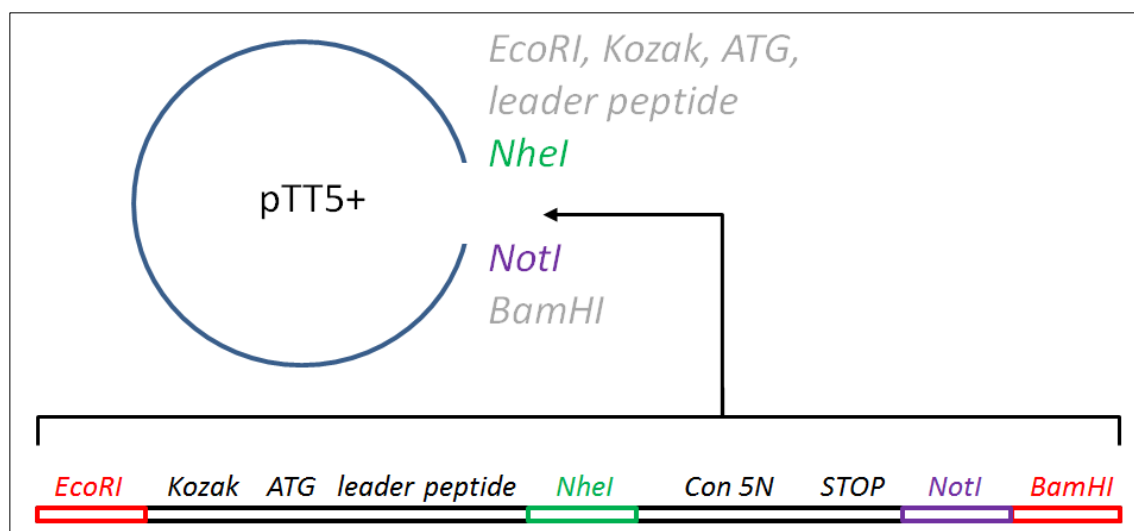


Figure 3-2: pTT5+ vector with removed MCS and inserted synthesized DNA fragment

The online tool *Double Digest Finder* from New England BioLabs was used to determine the best suitable buffer for the double digest, which was NEBuffer 2.1 (New England BioLabs).

The final volumes for the digestion are listed in Table 3-5 as well as the distinct recognition sequences of the enzymes used (Table 3-6).

Table 3-5: Restriction enzyme digest of vector

Component	Added volume
<i>NheI</i> [10 U/mL]	0.5 µL
<i>NotI</i> [10 U/mL]	0.5 µL
10x NEBuffer 2.1	5 µL
100x BSA	0.5 µL
pTT5+Con 5N plasmid	5 ug
dH ₂ O	To final volume (50 µL)

Table 3-6: Restriction enzymes used and their recognition sequences

Restriction enzyme	Recognition sequence
<i>NheI</i>	5' – G/CTAGC – 3'
<i>NotI</i>	5' – GC/GGCCGC – 3'

The digest mixture was incubated at 37 °C (ThermoMixer from Eppendorf) for 2 hours. To separate the resulting linearized vector from the removed insert, a preparative agarose gel electrophoresis was performed. The ingredients listed in Table 3-7 were melted in a microwave and 0.0001 % SYBR® Safe DNA gel stain (Invitrogen) was added before gels were casted.

Table 3-7: Composition of agarose gel for electrophoresis

Agarose (Biozyme)	1 %	1,5 g in 150 g
50x TAE	2 %	3 g in 150 g
dH ₂ O		To final volume

The 50 µL samples were mixed with 10 µL of *6X DNA Loading Dye* (Thermo Scientific) and then loaded onto the agarose gel. Additionally, 30 µL of *Gene Ruler 1kb DNA ladder* (Thermo Scientific) and *2 log DNA Ladder* (New England BioLabs) was loaded. Parameters for electrophoresis were set to 80 V and the gel ran for about 70 minutes.

The resulting bands were first checked under UV light using the *Gene Flash* apparatus for running at the expected size and then the linearized vector was excised from the gel. The DNA was purified using the *illustra GFX PCR DNA and Gel Band Purification Kit* from GE healthcare following the attached protocol. For elution 30 µL of the supplied buffer type 6 was used. The resulting DNA concentration was determined using NanoDrop 100 from Thermo Scientific.

3.1.2.2 Insert DNA amplification by PCR

The inserts were amplified by polymerase chain reaction (PCR). This molecular technique allows the amplification of DNA sequences by generating a high number of copies. PCR is based on thermal cycling, which includes repeated heating and cooling steps causing the denaturation of DNA, the annealing of primers to complementary regions and finally the enzymatic replication occurring in a polymerase chain reaction.

For the amplification a template was necessary as well as the respective primer pairs (as described above). The PCR was performed using the *Q5® High-Fidelity 2X Master Mix* from New England BioLabs. This mix already comprised the high fidelity DNA polymerase, deoxyribose nucleotides (dNTPs), Mg²⁺ as well as an optimized buffer. Only the specific template and primers had to be added.

Table 3-8: Insert preparation: components and concentrations used

Components	Final concentrations
Q5 2X Master Mix	1x
forward primer [20 µM]	1 µM
reverse primer [20 µM]	1 µM
template DNA	2 ng/µL
dH ₂ O	to 50 µL

The PCR mixtures were prepared in ultra-thin wall PCR tubes and then transferred to a Thermo Cycler (Thermo Scientific). The individual thermal steps used are summarized in Table 3-9.

Table 3-9: Thermo cycler settings for PCR: *touchdown* programme used for insert amplification.

Step		Temperature	Time (mm:ss)	Procedure
1	Initial Denaturation	98 °C	00:30	Initial Denaturation
2	10 Cycles: Starting at 72 °C annealing T was reduced by 1°C in each cycle	98 °C	00:10	Denaturation
3		72 °C ↓ 62 °C	00:20	Annealing
4		72 °C	02:20 (20-30 s/kb)	Extension
5	20 Cycles: annealing T is kept constant at 62 °C	98 °C	00:10	Denaturation
6		62 °C	00:20	Annealing
7		72 °C	02:20	Extension
8	Final extension	72 °C	10:00	Final Extension
9	Hold	4 °C	10:00	Cooling

3.1.2.3 Purification of DNA fragments from agarose gel

After PCR, the inserts were applied onto a 1 % agarose gel supplemented with SYBR Safe for preparative electrophoresis. As described in 3.1.2.1, DNA fragments were detected and purified in the same way (GFX purification).

3.1.2.4 Ligation

As mentioned before the *Gibson Assembly Cloning Kit* was used for the ligation of the linearized vector and the desired insert.

Gibson Assembly combines three different enzymatic reactions: An exonuclease creates single-stranded 3' overhangs that facilitate the annealing of the overlapping complementary regions. A DNA polymerase fills in gaps within each annealed fragment and a DNA ligase seals nicks in the assembled DNA. All these reactions can be combined to one isothermal single-tube reaction due to the Gibson Assembly Master Mix with the end result of a double-stranded fully sealed DNA molecule (schematic depiction in Figure 3-3).

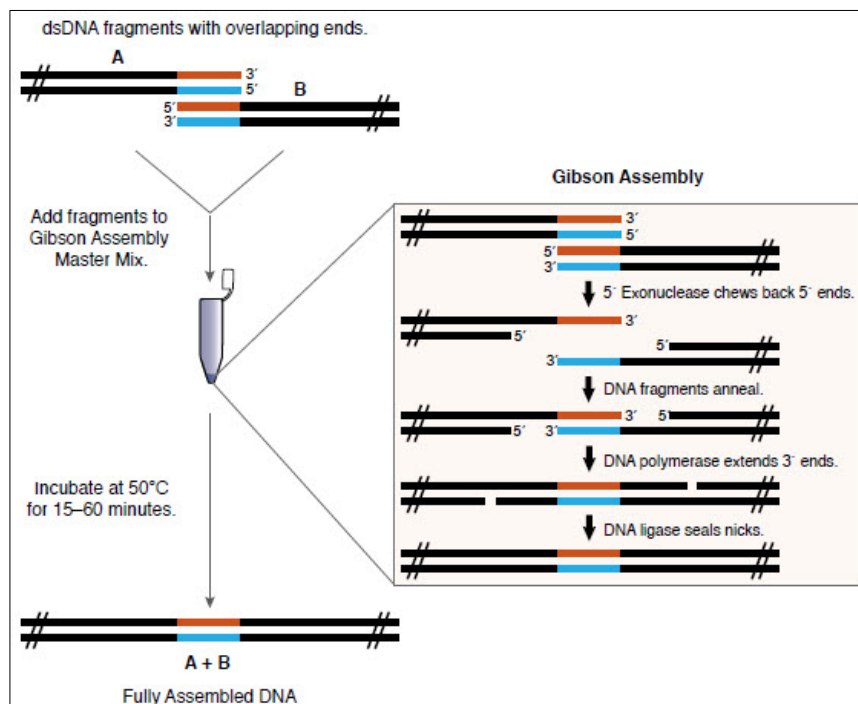


Figure 3-3: Gibson Assembly: scheme from homepage <https://www.neb.com/products/e5510-gibson-assembly-cloning-kit>

Based on the protocol 50–100 ng of vector should be ligated with a 2–3 fold molar excess of insert(s). To take the size of the vector and insert into consideration the amount of DNA in nanograms needs to be converted to picomoles. For this conversion the following equation was used:

$$pmol = \frac{ng * 1000}{bp * 650}$$

In total, vector and insert DNA should be in the range of 0.02 – 0.5 pmol. Here 0.1 pmol was chosen as total amount and therefore 0.025 pmol of the vector and 0.075 pmol of the distinct inserts were used.

The numbers of base pairs were determined for each construct using the programme *Gene Runner* and are summarized in Table 3-10.

Table 3-10: hsPxd01 variants and corresponding amount of base pairs

Variant	Base pairs
Con 1N	4401
Con 1N_helix	4173
Con 2N	3720
Con 3N	3882
Con 4N_helix	3516
Con 5N_cys	2142
Con 5N_helix	2391
Con 5N_VWC	2598

Table 3-11: Gibson Assembly mixture

Components	Final concentration
2X Gibson Assembly Master Mix	1x
Vector	0.025 pmol
Insert	0.075 pmol
dH ₂ O	To final volume
Final volume	20 µL

The Eppendorf tubes containing the mixtures (composition see Table 3-11) were incubated at 50 °C for 15 minutes and then transformed.

3.1.3 Transformation of chemical competent *E. coli* cells

During the incubation of the ligation step, competent cells were thawed on ice (NEB® 5-alpha Competent *E. coli* (High Efficiency) from New England BioLabs). 2 µL of the Gibson Assembly mixture were added to these competent cells. The tube was incubated on ice for 30 minutes and then heat shocked. This was done by placing the tubes in a 42 °C water bath for exactly 30 seconds. Following this, cells were placed on ice again for 5 minutes before 950 µL of room temperature SOC-medium (New England BioLabs) was added. This mixture was then incubated at 37 °C while shaking for 60 minutes.

A 100 and a 300 µL aliquot of transformed cells in SOC medium were then plated onto LB_{amp} agar plates (preparation see Table 3-12), respectively, and incubated overnight at 37 °C for selection due to resistance against ampicillin.

All the transformations of constructs were monitored by carrying out positive and negative controls.

Table 3-12: Preparation of LB medium and LB agar: components were dissolved in dH₂O and autoclaved at 121 °C and 2 bar for 25 minutes. Ampicillin was added directly before usage.

Medium	Composition
LB-medium	1 % (w/v) peptone 0.5 % (w/v) yeast extract 1 % (w/v) NaCl
LB-agar	1 % (w/v) peptone 0.5 % (w/v) yeast extract 1 % (w/v) NaCl 1.5 % (w/v) agar
Ampicillin	0.1 mg/mL final concentration

3.1.4 Screening for positive clones

From the resulting plates a number of clones were picked for screening. Each colony picked was subjected to PCR Screening (3.1.4.1), transferred to an LB_{amp}-masterplate and was further used to inoculate an overnight culture for plasmid preparation (3.1.4.2). The plasmid preparation (mini prep) was used for a restriction enzyme digest to screen for the correct insert size and subsequent DNA sequencing. Cryo stocks (3.1.5) of sequenced clones with correct insert DNA sequence were prepared.

3.1.4.1 *PCR Screening of transformed E. coli cells*

For evaluating which colonies from the LB_{amp}-plate carried the correct insert in the vector a PCR screening was performed. Therefore variable numbers of colonies (5 – 10) were picked sterilely with a toothpick and added to a prepared mixture of the following components (Table 3-13):

Table 3-13: Composition of mixture prepared for PCR screening

Components	Final concentration
Q5 2X Master Mix	1x
pTT5 forward primer	0.5 μ M
pTT5 reverse primer	0.5 μ M
dH ₂ O	To final volume
Final volume	20 μL

Then again a PCR was performed as described in Table 3-9 with the difference that the initial denaturation step was elongated to 10 minutes to allow complete denaturation of cells and release of plasmid DNA to serve as PCR template. Once PCR was finished 4 μ L of a 6x loading dye was added to the PCR mixture which was then applied to an analytical 1 % agarose gel supplemented with SYBR Safe for electrophoresis. The results were again visualized by UV-light, as already described above.

3.1.4.2 Mini Prep: Plasmid DNA preparation of transformed *E. coli* cells

As mentioned before (3.1.4), an overnight culture was prepared. Therefore 3 mL LB_{amp} medium were inoculated with a sterilely picked colony and incubated at 37 °C for 16 hours while shaking (180 rpm). The next day cells were harvested by centrifugation (microcentrifuge: Thermo Sigma 1-15PK, 1 min, 16000 g) and plasmid DNA was purified following the manual of *illustra plasmidPrep Mini Spin Kit* from GE Healthcare life Sciences. The principle will be described in more detail in chapter 3.1.6. The obtained amount of plasmid DNA was measured at 260 nm using NanoDrop 100 (Thermo Scientific).

The plasmid DNA could be further used for a restriction enzyme screening. A double digest with the restriction enzymes *EcoRI* and *SspI* (New England BioLabs) was performed.

Table 3-14: Restriction enzymes and respective recognition sequence

Restriction enzyme	Recognition sequence
<i>EcoRI</i>	5' – G/AATTC – 3'
<i>SspI</i>	5' – AAT/ATT – 3'

For this double digest again the appropriate buffer was determined using the Double Digest Finder Tool. For *EcoRI*-HF and *SspI* the *10x Buffer 2.1* (NewEngland Biolabs) is suitable. The final composition of the digest is listed in Table 3-15.

Table 3-15: composition of the double digest of plasmid DNA

Component	Added volume
<i>EcoRI</i> -HF [20 U/mL]	0.5 µL
<i>SspI</i> [5 U/mL]	0.5 µL
10x NEBuffer 2.1	2 µL
100x BSA	0.2 µL
Plasmid DNA [100 ng/µL]	10 µL
dH ₂ O	To final volume (20 µL)

After 2 hours of incubation at 37 °C the digest was applied to an analytical agarose gel and the obtained fragments were separated by electrophoresis. Two bands should appear, as both enzymes are single cutters. The program gene runner was used to determine the expected sizes of the fragments for each construct.

3.1.4.3 Plasmid DNA sequencing

Sequencing was used to verify and validate the final plasmid DNA products. Depending on the length of the respective constructs different primers were used for sequencing so that in the end the whole length of the distinct construct is covered.

Table 3-16: Primers used for DNA sequencing

Name		Sequence 5' → 3'
pTT5 fw	forward	AAAAACGAGGAGGATTTGATATTCACCTGG
606fw		CCTGTGGTTGGCGGATTG
2038fw		ATCTTTGAACGGACATTGC
seq2_fw		CATCGTGCAGCGGTCC
seq3_fw		GGCTCTTCCTCAGTTCAC
pTT5 rev	reverse	CCGAGGGATCTCGACCAAATGATTTGC
3119 rev		ATCTTCGGGAGCCAGTGC
4130 rev		GTGCTGAAGGCTGAGGTGC

Every vial sent for sequencing contained:

- 400 – 600 ng of plasmid DNA
- Distinct primer with a final concentration of about 1.4 μ M
- dH₂O to fill up to a total volume of 14 μ L

Depending on the results of previous screening experiments like PCR Screening or double-digest of plasmid DNA, the purified plasmid DNA of either 1 or 2 clones of one construct was submitted for DNA sequencing to Microsynth AG where sequencing was done overnight using Sanger sequencing technology. The obtained results were evaluated by the sequencing program *seqman*.

3.1.5 Preparation of cryo cultures

To preserve the validated clones of each construct, cryo cultures were generated. Therefore, 500 µL of an overnight culture were mixed with 500 µL glycerol (30 %) under sterile conditions in a cryo tube. Due to the addition of this cryo protective agent, the clones can be stored at -80 °C for a long period of time.

3.1.6 Plasmid isolation from *E. coli* cultures

For transfection and large scale expression of the distinct constructs larger amounts of plasmid DNA were required than for example for DNA sequencing as explained in chapter 3.1.4.2 and 3.1.4.3. Depending on the amount needed either mini, midi, mega or giga prep kits are available.

All of these kits have in common that the purification is based on selective alkaline denaturation of high molecular weight chromosomal DNA. At the same time covalently circular DNA remains double-stranded (Birnboim and Doly, 1979).

For the transient transfection in relatively large scale as described in the following chapters, purifications with *Mega Prep Kits* were necessary in order to gain enough DNA. Therefore the *Plamid Plus Mega Kit* from QIAGEN was used. 1 litre of LB_{amp} medium was inoculated with the respective cryo stock of the distinct hsPxd01 variant and grown at 37 °C overnight while shaking with 180 rpm. After at least 16 hours the cells were harvested by centrifugation at 6000 g for 15 minutes. The resulting pelleted bacteria were further treated following the attached *Purification Handbook* from QIAGEN.

The concentration as well as the purity of the resulting plasmid DNA was then measured using NanoDrop 100 from ThermoScientific. The isolated DNA was finally stored at -20 °C until usage.

3.2 Protein production

3.2.1 Human embryonic kidney cells in cell culture

The transient transfection of human embryonic kidney cells is a well-established cell system for the expression of mammalian proteins as these cells are able to perform posttranslational modifications like glycosylation or disulphide bridge formation. These cells can be cultivated in serum-free media which is a great advantage for purification. The cell line used (HEK 293-6E) stably expresses the Epstein-Barr virus nuclear antigen 1.

Expression of truncated human peroxidase variants in HEK 293-6E cells was established and optimized previously and the protocol was applied in this work here. For cultivation polycarbonate flasks with vented caps (Corning) were used.

3.2.2 Sterile technique

Minimizing the potential risks of contaminations is a crucial factor in cell culture handling. Therefore some preventive arrangements were kept strictly whenever working with cells or components which came in contact with cells. Every action was performed in a laminar flow hood, which was powered up at least 10 minutes before working by turning on the fan as well as a UV-lamp. Before cell culture material was taken into the bench, the surface was cleaned with 70% ethanol. All bottles, vessels, flasks or other materials were also disinfected by ethanol before placing in the hood. After finishing work in the laminar flow bank, it was cleaned again before turning off.

3.2.3 Preparation of culture media and additives

For the cultivation of HEK cells the FreeStyle™ F17 Expression Medium (Gibco) was used. Some supplements (see Table 3-17) had to be added before use and for stability reasons the medium was used up within 7 days.

Table 3-17: Supplements for F17 cell culture medium.

Supplement	Necessary for	Stock solution	Final concentration
Pluronic F68 (ThermoFisher Scientific)	Stabilisation of cell membranes, protection against shearing	10 %	0.1 %
L-glutamine (Gibco)	Essential amino acid, acts as a source of energy	200 mM	4 mM
Geneticin (Biochrom)	Antibiotic, selection marker	50 mg/mL	50 µg/mL

Moreover a 20 % peptone stock solution was prepared. Therefore 100 g of TN1 (AppliChem) was dissolved in cell culture grade water (Life Technologies) to give a total volume of 500 mL. This solution was filtered by attaching a SteriTop Filter Unit (Merck) on a sterile Schott bottle and filtering the solution by applying vacuum. This system was operated in the Laminar Flow bench.

For heme incorporation, porcine hematin (Sigma Aldrich) was added in the form of a 2 mg/mL solution with a final concentration of 5 µg/mL. The substance was dissolved in 0.1 M NaOH. This solution was prepared freshly before use and filtered using a syringe (Omnifix) and a syringe filter (0.22 µm, Roth) to make the solution sterile before adding to the cells.

3.2.4 Thawing of HEK 293-6E cells

Working cell banks of HEK cells are stored in liquid nitrogen at -196 °C. Before thawing cells, 16 mL of F17 expression medium, complemented with pluronic and L-glutamine but without geneticin, were adapted to the cultivation conditions. The medium was incubated in a 25 mL shaker flask and put in the incubator on the orbital shaker (120 rpm, 37 °C, 5 % CO₂, 70 % humidity). Cells were thawed rapidly (hand thawing) to prevent cell damage and transferred to the pre-conditioned medium. The cryo cultures contained approximately 7.5*10⁶ cells/mL which resulted in cell density of about

0.47*10⁶ cells/mL. Dilution of the cells was first performed when they reached a density of 1*10⁶ cells/mL. Geneticin was added to the medium as soon as the cells established a cell doubling time of 24 hours.

3.2.5 Maintenance of HEK cells

HEK cells have to be passaged every 2-3 days due to different issues: The media supplements L-glutamine and geneticin are not stable in the expression medium for longer than 5 days when incubated at 37 °C. Due to the rapid growth of cells in suspensions, the nutrients are metabolized rapidly and the accumulation of toxic substances is a potential danger to the cells. By passaging the cell density can be kept in the optimal range of 0.25 * 10⁶ to 2 * 10⁶ cells/mL.

Therefore the density and the viability was determined three times a week (Monday, Wednesday and Friday). This was done by taking an aliquot of 50 µL from the flask, diluting it with 200 µL 1x PBS (Gibco) and dying the cells with 50 µL Trypan blue solution (Gibco). Dead cells would be coloured blue as Trypan blue is able to enter the cell due to disrupted membranes. The diluted cell solution was then evaluated under a microscope and cells were counted using a hemocytometer.

3.2.6 Transient Transfection of HEK 293-6E cells

All constructs were expressed via transient transfection. In comparison to stable transfection, the DNA introduced is not integrated into the genome but only present for a limited time. The gene of interest was brought into the cells with the pTT5+ vector. Further a cationic polymer is needed to shuttle the plasmid DNA into the cells. Polyethyleneimine (PEI) was used because of its ability to condense DNA into positively charged particles. These are able to bind to the cell surface and are finally endocytosed by the cell. Finally DNA is released into the cytoplasm and shuttled into the nucleus for transcription (Longo *et al.*, 2013).

Transient expression was performed following an established protocol with a 7 day cycle. One week prior to transfection cells were passaged and thereby the volume of cell

suspension was increased to have a final transfection volume of about 2-3 L, depending on cell densities. Two days before transfection cells were diluted for the last time to a density of 0.45×10^6 cells/mL. Directly before transfection the density was checked again to be between $1.5 - 2.0 \times 10^6$ cells/mL (ideal: 1.7×10^6 cells/mL).

For transfection a mixture of PEI and the DNA which should be introduced had to be prepared. The two components were each diluted in expression medium (the volume used therefore was 5 % of the total cell suspension volume). So for example for 3 L of cell culture PEI and the DNA were each added to 150 mL F17 medium in separate flasks. PEI was added to one of the 150 mL F17 medium flasks with a final concentration of 2 $\mu\text{g/mL}$ in the flask, using a stock solution with a concentration of 1 mg/mL. DNA was gained from plasmid purification and thus concentrations were varying. Thus to the other flask of 150 mL of F17 medium a total amount of DNA was added to result in a final concentration of 1 $\mu\text{g/mL}$ of the total cell suspension volume (3mg added for 3L transfection). The two solutions containing PEI and DNA were then poured together and incubated for 20 minutes. Then the appropriate amounts were added to the flasks (10 % of the flask volume).

Four hours post transfection the hematin solution (see 3.2.3) was added to the flasks to achieve a final concentration of 5 $\mu\text{g/mL}$.

Two days after transfection peptone was added to the cells. The TN1 solution was added to gain a final concentration of 0.5 % per flask.

3.2.7 Harvest of cells

The supernatants containing the soluble protein were harvested three days after the addition of TN1. This was done by centrifugation at 16900 g for 30 min (Thermo Sorvall RC6 Plus). The resulting pellet was discarded (autoclaved) and the supernatant was filtrated using a vacuum filtration system and a durapore membrane filter with 0.45 μm pore size (Merck). The filtrated supernatant was filled into SCHOTT bottles and frozen at $-30\text{ }^{\circ}\text{C}$ until it was used for protein purification.

3.3 Protein purification

3.3.1 Diafiltration

The harvested supernatant was thawed either slowly at 4 °C or in a waterbath at 25 °C. Then the sample was concentrated to a final volume of about 100 mL (20 to 50-fold depending on starting volume) and exchanged into a 100 mM phosphate buffer, pH 7.4 (composition see Table 3-18). Both steps were done using a Millipore labscale diafiltration system equipped with 3 Pellicon® membranes (MWCO 30 kDa) from Merck Millipore.

Table 3-18: Buffer for diafiltration

Buffer (4 L, pH 7.4)	9.12 g KH_2PO_4 86.92 g $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$
-----------------------------	--

3.3.2 Immobilized metal ion affinity chromatography (IMAC)

The diafiltrated sample was adjusted to the right conditions for IMAC, which corresponded to the concentration of the equilibration Buffer A (see Table 3-18). The volume was measured and appropriate amounts of a 2 M imidazole solution (final concentration of 5 mM) as well as a 4 M NaCl solution (final concentration 1 M) were added. The pH had to be adjusted to 7.4 using a KOH solution. Finally the sample was filtered through a 0.45 µm durapore membrane filter using a vacuum pump.

For purification the ÄKTA purifier FPLC-system (GE-Healthcare) was used. As the different constructs all include an N-terminal hexa-histidine-tag, they all bind to nickel.

Two HisTrap FF crude 5 mL columns (GE healthcare Life Sciences) were coupled, loaded with a Ni-solution and connected to the ÄKTA system. First they were equilibrated with Buffer A and then the sample was applied to them. Washing with Buffer A after sample loading should remove all non-specifically bound materials from the column. The absorbance at 280 nm and 410 nm was monitored and as soon as it

dropped to its initial value before sample application elution was started. Therefore a gradient slowly increasing imidazole concentration by mixing Buffer A and Buffer B was chosen. The eluted samples were collected in 2 mL fractions using a fraction collector.

Settings for flow rates and gradients were slightly varied for each construct. Standard settings for purification of Con 4N are listed in Table 3-20.

Table 3-19: Buffers and solutions for IMAC purification

Buffer A (pH 7.4)	100 mM phosphate buffer, pH 7.4 1 M NaCl 5 mM imidazole
Buffer B (pH 7.4)	20 mM phosphate buffer, pH 7.4 500 mM NaCl 500 mM imidazole
Ni solution	0.1 M NiCl ₂
Stripping solution	1x PBS 50 mM EDTA
Regeneration	1 M NaOH
Neutralization	200 mM phosphate buffer, pH 7.4
Storage solution	20% (v/v) ethanol

Table 3-20: Settings used for purification with ÄKTA system: standard settings for purification of Con 4N

Settings for purification	
Max. pressure	0.3 MPa
Flow rate for equilibration	3 mL/min
Flow rate for sample application	1.5 mL/min
Flow rate for washing	3 mL/min
Flow rate for gradient 1	2 mL/min
Flow rate for gradient 2 (elution)	1 mL/min
Gradient 1	0 – 8 % B in 10 min
Gradient 2	8 – 70 % B in 50 min

Fractions eluting during the second gradient should contain the desired protein and were checked using UV-Vis spectra. Thereby the absorption at 280 nm as well as the typical heme Soret peak at about 410 nm (varying for each construct) was determined. This will be explained in detail in chapter 3.4.1.

SDS-PAGE as well as Western Blots were performed in addition (see 3.4.2 – 3.4.4), as the UV-Vis spectra alone would not give enough information about the purity of the fractions. Also before the Western Blots were developed they were tested for covalently bound heme using enhanced chemiluminescence.

The results of spectra, SDS-PAGE, ECL and Western Blots were then used to choose fractions which were pooled and used for the following steps.

3.3.3 Concentrating

The pooled fractions, containing the protein of interest, further had to be concentrated. So the volume could also be reduced. Therefore the pool was transferred to a SnakeSkin® Dialysis Tubing with a molecular cut-off of 10 kDa (Thermo Scientific). For concentration the tube was covered with polyethylene glycol (PEG) with a molecular weight of 20 kDa which was available as white flakes (Roth). This water-soluble polymer is highly hygroscopic and therefore concentrates the sample in the tubing. Depending on the starting volume and protein concentration it took 2 to 4 hours to reduce the sample volume to about 4 to 8 mL. In the meantime polyethylene glycol was renewed from time to time. Finally the dialysis tube was cleared from polyethylene glycol by washing with dH₂O.

The sample in the dialysis tubing could be used directly for subsequent dialysis.

3.3.4 Dialysis

The concentrated sample was further dialysed to remove salt and imidazole which was present after purification. Dialysis took place over-night at 4 °C. Therefore the sample which was already located in the SnakeSkin Dialysis Tubing was placed in the exchange

buffer (Table 3-21) and dialyzed while stirring. The buffer was renewed once after approximately 3-4 hours, which resulted in two subsequent dialysis steps with a sample to buffer ratio of at least 1:100 in each of the steps.

Table 3-21: Exchange buffer for dialysis

Buffer (2 L, pH 7.4)	4.56 g KH ₂ PO ₄
	43.46 g Na ₂ HPO ₄ ·7H ₂ O

After dialysis the sample was removed from the tube and the final concentration was determined by spectroscopy. The sample was aliquoted in appropriate amounts in 1.5 mL Eppendorf tubes and either used immediately for characterisation or was temporarily frozen at -30 °C.

3.4 Protein characterization

3.4.1 UV-Vis spectroscopy

To determine the spectral properties of the different variants absorption spectra were measured by filling the sample into a 10 mm quartz cuvette and using an Agilent-8453 UV-Visible spectrophotometer. Thereby a spectral wavelength scan in the range of 200 – 800 nm was performed and the absorptions at 280 nm as well as at 410 nm were displayed. At 280 nm aromatic amino acids absorb generating the overall protein peak. The characteristic heme Soret peak is usually observed at a wavelength between 400 and 430 nm, depending on the heme protein, due to heme coordination and/or modification. The ratio between these two values is the so-called *Reinheitszahl (RZ)* or *purity number*. It can be determined as followed:

$$RZ = \frac{Abs\ 410\ nm}{Abs\ 280\ nm}$$

The concentration of a protein solution can be calculated by using the Lambert-Beer-Law:

$$A = \varepsilon * c * d$$

A indicates the absorbance [AU], **ε** stands for the molar extinction coefficient [$M^{-1} \cdot cm^{-1}$], **c** is the concentration [$mol \cdot L^{-1}$] and **d** the diameter of the cuvette [cm].

After calculating the concentration the total amount of the variant **m [g]** could be determined as well. Additionally to the concentration **c [mol/L]** the molecular mass **M [g/mol]** and the volume **V [L]** must be known. Calculation then could follow the equation:

$$m = c * M * V$$

Spectra were always measured with the protein being diluted in the distinct buffer. Usually this was a 100 mM phosphate buffer, pH 7.4. The corresponding blank was collected with the same buffer.

3.4.2 SDS-PAGE

Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis (SDS-PAGE) is probably the most commonly used method to separate protein solutions based on to their size and in further consequence the molecular mass of the respective components can be determined. The polyacrylamide gel, to which samples are applied, is an inert homogenous carrier and SDS is an anionic detergent, forming a complex with the protein and thereby inducing a negative charge correlating with the length of the peptide chain. The established protein-SDS complex moves with a distinct velocity which is directly proportional to the logarithm of its molecular mass. Mobility of charged molecules in an applied electric field is therefore depending on charge, shape and size. The movement is further determined by ionic strength, viscosity and medium temperature and so it is essential to track the path with a marker of known protein masses. To visualize the resulting bands, the gel has to be stained for example with Coomassie staining solution which is a relatively quick and easy solution.

For sample preparation 15 µL of the protein solution was incubated with 5 µL of a 4x loading buffer (NuPAGE LDS Sample buffer from Life Technologies) at 70 °C for at least 10 minutes. This causes the protein to denature. For SDS-PAGE under reducing conditions, DTT was added to the assay (10x DTT stock at 0.5 M). About 19 µL sample were then applied to the *NuPAGE Novex 4-12% Bis-Tris Protein Gel 1.0mm x 15 well* (Life Technologies) placed in an SDS-PAGE chamber and filled with 1x MOPS buffer (dilution of NuPAGE MOPS SDS Running Buffer 20x from Life Technologies). Additionally 12 µL of a protein standard was applied.

The gel then ran for about 45 minutes with voltage set to 200 V. As soon as the running front had reached the end of the gel, it was removed and transferred to a plastic container for staining. The gel was therefore covered with staining solution (Table 3-21) and incubated at room temperature while shaking. After approximately 45 min the gel was rinsed with dH₂O and incubated in destaining solution (Table 3-22) under the same conditions until the bands were clearly visible and the background was de-coloured.

Table 3-22: Staining and destaining solution: these solutions were prepared for Coomassie staining and destaining of SDS-PAGE gels

Staining solution	Destaining solution
40 % methanol	40 % methanol
10 % acetic acid	10 % acetic acid
50 % dH ₂ O	50 % dH ₂ O
0.1 % Coomassie Blue R-250	

3.4.3 Western Blot

Western Blotting is a common method to transfer protein bands from an SDS-PAGE to a carrier membrane and detect them with specific antibodies to further visualize only the desired protein. Transfer is done by pulling the proteins from the gel due to electric current.

By using the iBlot Dry Blotting System this method can be performed very easily. The pre-run gel was placed on an anode stack in the device, covered with a pre-soaked (dH₂O) filter paper and finally covered with a cathode stack. All these steps were carried out according to the attached protocol. The parameters were set to 20 V for 7 minutes. After this the gel was removed and the nitrocellulose membrane was either used for Enhanced Chemiluminescence (3.4.4) or immediately incubated with the respective antibodies (see Table 3-23). Therefore it was transferred into a plastic container and covered with blocking buffer (2 % (w/v) BSA (Sigma) and 0.02 % Tween in 1x PBS buffer) and incubated while shaking at room temperature. After at least 30 minutes the primary antibody was added (1:2000 dilution) and incubated for at least 4 hours but usually overnight, at 4 °C while shaking. After this the solution was removed and the membrane was washed with PBST (0.02 % Tween in 1x PBS buffer) three to four times for 5 to 10 minutes each. The same procedure was repeated with the second antibody.

For visualizing the protein of interest the membrane was covered with 5 mL of an alkaline phosphatase buffer to which 33 µL NBT and 16.5 µL BCIP (Promega) were added. The membrane was incubated at room temperature until bands were visible (approximately 5 minutes). The reaction was then stopped by rinsing with dH₂O and the developed blot could be scanned.

Table 3-23: Antibodies used for Western Blot and respective dilution in binding buffer

Antibody 1	Penta His Antibody BSA-free (100 µg) Mouse monoclonal IgG ₁	QIAGEN	1 : 2000
Antibody 2	Anti-Mouse IgG (Fc-specific)-Alkaline Phosphatase, antibody produced in goat Affinity isolated antibody, adsorbed with human IgG	SIGMA	1 : 50000

3.4.4 Detection of heme by enhanced chemiluminescence

The detection of covalently linked heme is especially important when a protein with covalent heme to protein linkages is expressed for the first time to confirm correct heme insertion. These covalent linkages can be revealed using enhanced chemiluminescence (ECL). Therefore an SDS-PAGE was performed and the resulting gel was blotted (see 3.4.2 – 3.4.3). Due to the fact that non-covalently bound heme is lost during SDS-PAGE only the covalently linked heme will be detected.

The blot was covered with 1.5 mL of a mixture of equal amounts of two ECL solutions (Clarity® Western ECL Substrate Peroxide Solution and Luminal/enhancer solution from BioRad). The contained luminol reacts with the heme and as a further consequence light is emitted. The blot was placed into the ChemiDoc® MP System (BioRad) where measurements took place. Afterwards the membrane could be used for further experiments like Western Blotting, starting at the incubation step with blocking buffer.

3.4.5 Circular dichroism spectroscopy

Circular dichroism makes use of the fact that left and right circularly polarized light is absorbed by chiral molecules to a different extent. Chirality of substances can either be derived from carbon atoms with four different substituents, disulphide bonds, being intrinsically chiral, or secondary structural elements. Furthermore, chirality is observed when aromatic amino acids are linked to asymmetric centres within biomolecules, e.g. proteins.

The different constructs were diluted to a concentration of 10 μ M in 100 mM phosphate buffer, pH 7.4. CD spectra measuring the circular dichroism [mdeg] at different wavelengths were then recorded.

3.4.6 Redox potential

Covalently linked heme influences the heme distortion and therefore has an impact on the redox properties. Determination of the standard reduction potential (E°) of the Fe(III)/Fe(II) couple of the different constructs was carried out in Modena, Italy, by the working group of G. Battistuzzi. 2.5 mg of the respective samples were sent frozen at a concentration of approximately 1 mg/mL and measured using a homemade optically transparent thin layer spectroelectrochemical cell performing spectroelectrochemical titrations. To determine the E° of a redox couple, a series of potentials are sequentially applied. At each distinct potential, the ratio of the concentrations of oxidized to reduced forms, of the redox couple is given by the Nernst equation:

$$E = E^\circ + \frac{RT}{nF} \ln \frac{[\text{Ox}]}{[\text{Red}]}$$

E applied potential

E° formal potential of the redox couple Ox/Red

n number of electrons transferred

R molar gas constant (8.31 J.mol⁻¹.K⁻¹)

F Faraday's constant (96.48 C/equiv.)

Based on the known Soret maxima of the ferric ($A_{\text{ox}}^{\text{max}}$) and ferrous ($A_{\text{red}}^{\text{max}}$) states of the constructs and the respective absorbance values of defined applied potentials (i.e. A_{ox} , A_{red}), the following equation which represents the x-axis is obtained:

$$\log \frac{A_{\text{red}}^{\text{max}} - A_{\text{red}}}{A_{\text{ox}}^{\text{max}} - A_{\text{ox}}}$$

The intercept of the line of this plot then represents the value of E° . The slope is related to n and refers to the number of exchanged electrons.

3.4.7 Activity measurements

3.4.7.1 Peroxidase activity assays

Enzymes with peroxidase activity can be tested quite easily with well-established assays. Substrates like the chemical compound ABTS or guaiacol, are oxidized and thereby yield coloured products. These reactions, resulting in an increasing absorbance, can be measured photometrically (HITACHI UV-3900).

ABTS assay

ABTS, short for 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid), gets oxidized by the enzyme while hydrogen peroxide is reduced. The reaction can be monitored by measuring the increasing absorbance at 414 nm ($\epsilon_{\text{ABTS ox } 414 \text{ nm}} = 31100 \text{ M}^{-1}\text{cm}^{-1}$). Table 3-23 shows the composition of the assays. The first measurements were carried out with Con 4N only. The pH of the buffer was varied to determine the pH-dependence of the overall reaction and the optimal pH range to work in. In this pH-region also other constructs were measured and compared to each other. Reactions were always started by addition of the enzyme.

Table 3-24: Composition of ABTS assay

Component	Final concentration
Enzyme	50 nM
Hydrogen peroxide	200 μM
ABTS	1 mM
Phosphate buffer	100 mM
dH ₂ O	To final volume

Guaiacol assay

Oxidation of guaiacol by a peroxidase results in an increase in absorbance at 470 nm. The composition of the assays is described in Table 3-25. Again, the reaction was

started by the addition of the enzyme and followed photometrically at 470 nm ($\epsilon_{\text{Guaiacol ox}}_{470\text{nm}} = 26600 \text{ M}^{-1}\text{cm}^{-1}$).

Table 3-25: Composition of guaiacol assay

Component	Final concentration
Enzyme	50 nM
Hydrogen peroxide	200 μM
Guaiacol	1 mM
Buffer	100 mM
dH ₂ O	To final volume

3.4.7.2 Halogenation activity assays

The reaction of hsPxd01 variants with the halides bromide and thiocyanate was measured spectrophotometrically using the HITACHI UV-3900 Spectrophotometer.

Bromination assay

For determination of bromination activity an indirect assay was used. Therefore the halogenation of NADH was monitored. Human peroxidase 1 catalyses the oxidation of bromide to hypobromous acid which further reacts with NADH. As a result, bromohydrin is produced which can be detected at a wavelength of 275 nm. For evaluation an extinction coefficient of $\epsilon_{\text{Bromohydrin } 275 \text{ nm}} = 11800 \text{ M}^{-1}\text{cm}^{-1}$ was used (Soudi *et al.*, 2015).

The measured sample had a total volume of 800 μL and was composed of 100 nM enzyme, 20 mM phosphate buffer, pH 7.4 and 100 μM NADH (Sigma). Varying concentrations of sodium bromide (NaBr from Merck) solution and hydrogen peroxide (H_2O_2 from Sigma) were added. For assays with varying NaBr concentrations (range: 0 – 50 mM) H_2O_2 concentration was held constant at 50 μM . If H_2O_2 concentration was varied (range: 0 – 500 μM) NaBr was kept constant at 5 mM. All reactions were started by the addition of hydrogen peroxide and the initial reaction rates were used to determine K_M and v_{max} by non-linear regression.

Further, this assay was used to determine the pH dependence of the over-all reaction. Therefore the composition of the compounds was held constant and only the pH in phosphate buffer was varied. Due to restricted buffering capacity of phosphate buffer, a citrate buffer was used for pH 5.

Thiocyanate oxidation assay:

Another indirect method to determine substrate turnover represents the so-called TNB assay. Human peroxidase 1 catalyses the oxidation of thiocyanate to hypothiocyanate (OSCN^-) which is detected with the help of 5-thio-2-nitrobenzoic acid (TNB).

A TNB solution was prepared starting with a 1 mM DTNB- solution which is also called Ellman's reagent. Therefore 5,5'-dithiobis-(2-nitrobenzoic acid) (Sigma) was dissolved in 50 mM phosphate buffer, pH 7.4. By subsequent rising of the pH to 12 DTNB is cleaved into 2 units of TNB which can be monitored by the solution turning yellow. After a few minutes the maximum of colour development was reached and the pH could be adjusted to 7.4 again. This TNB solution could then be used to assay the generation of OSCN^- . Therefore the extent to which the yellow solution was bleached was measured. 2 mol of TNB are oxidized to 1 mol of DTNB, which is colourless again, by reacting with hypothiocyanate, resulting in a loss of absorbance at 412 nm. Table x shows the components of the TNB assay.

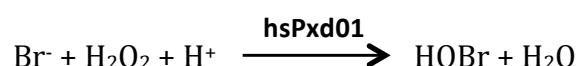
$$\epsilon_{\text{TNB } 412 \text{ nm}} = 14100 \text{ M}^{-1} \text{ cm}^{-1}$$

Table 3-26: Composition of TNB assay: thiocyanate was available in the form of KSCN (Sigma).

Component	Final concentration
Enzyme	100 nM
Hydrogen peroxide	200 μM
KSCN	100 mM
TNB	100 μM
Phosphate buffer, pH 7.4	100 mM
dH ₂ O	To final volume

3.4.8 Polarographical measurement of hydrogen peroxide consumption

In comparison to the indirect assays for halogenation activity as mentioned above, the turnover of substrates can also be measured directly using a so-called H₂O₂ electrode (Amperometric Biosensor Detector 3001 from Universal Sensors). This platinum electrode is covered with a hydrophilic membrane and has to be calibrated against a distinct H₂O₂ concentration. Further the applied potential at the respective pH value has to be adjusted (650 mV at pH 7.0). The loss of hydrogen peroxide concentration in a system containing the enzyme, H₂O₂ and a halide is thereby monitored polarographically. The depletion of hydrogen peroxide is explained by the ability of human peroxidase 1 to catalyse the oxidation of halides and pseudohalides using hydrogen peroxide.



Measurements were performed at physiological substrates concentrations (Table 3-26).

Table 3-27: Substrates used for polarographical measurements

Final concentration	Substrate	Available in form of	From
100 µM	Bromide	NaBr	SIGMA
	Thiocyanate	KSCN	SIGMA
	Urate	Uric acid	SIGMA
	Tyrosine	L-Tyrosine	SIGMA
	Ascorbic acid	L-Ascorbic acid	SIGMA
140 mM	Chloride	NaCl	SIGMA

The respective substrate and 100 mM phosphate buffer, pH 7, 100 µM H₂O₂ and 100 µM EDTA – were incubated in the measuring chamber while stirring. Then the reaction was started by the addition of 500 nM hsPxd01 construct 4N using a Hamilton syringe. Additionally a negative control was collected by adding no substrate to measure the hydrogen peroxide consumption by the enzyme.

3.4.9 Pre-steady-state kinetics

The transient-state or pre-steady-state describes the phase before a reaction reaches the steady-state and takes place in a time range of milliseconds. These reactions can be followed spectroscopically by stopped-flow spectroscopy. The apparatus for this technique consists of a mixing device, an observation cell, a stop syringe, a detection system and a record system (see Figure 3-4). The latter is capable of responding to rapid changes. The compounds (at least two – enzyme and substrate/ligand) are mixed rapidly and their pre-steady-state kinetics in solution are determined. The dead time of the apparatus, describing the time between mixing and the beginning of observation, has to be considered because during this interval no measurements are possible. After this dead time (which accounts usually for 1 millisecond) one can be sure that the reactants have reached the measuring cell.

Figure 3-4 depicts a simplified scheme of a conventional stopped-flow mixing mode. Two reactants (A and B) are prepared separately in two syringes. By operating the mechanical drive, both syringes are pushed simultaneously. The substances are mixed and further forced into the measuring cell. From there the mixture is forwarded to the stop syringe which shortly moves to cause a mechanical stop and thereby prevents further mixing. Additionally the detecting system gets activated.

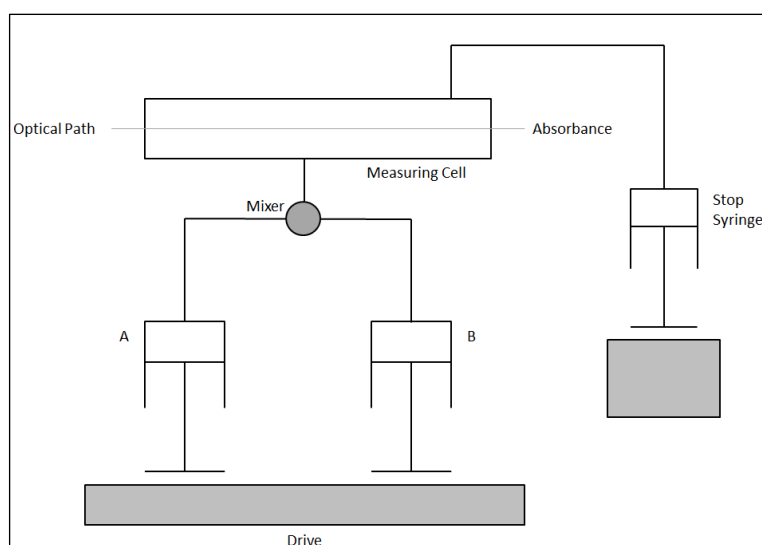


Figure 3-4: Scheme of stopped-flow mixing mode

Besides this just described conventional mode, it is also possible to operate stopped-flow mixing systems in sequential mode. In this case up to four reactants can be utilised to perform two reaction steps.

It is possible to perform different methods of detection with this apparatus. When using a photodiode array detector, spectra can be recorded. On the other hand single wavelengths can be measured by applying a photomultiplier.

3.4.9.1 Reaction with hydrogen peroxide – Compound I formation

The addition of hydrogen peroxide or other peroxides to the ferric enzyme causes the formation of Compound I which represents the first step in both the halogenation cycle and the peroxidase cycle.

In principle a peroxidase can be oxidized by H_2O_2 . The enzyme is converted from its resting state to a reactive form, called Compound I which has a high redox potential. Hydrogen peroxide is reduced to water. This Compound I then can oxidize a two-electron donor. A donor could be a halide, like bromide, or also chloride or iodide. So the halide is oxidized to its corresponding hypohalous acid and the enzyme is thereby reduced back to its ferric resting state.

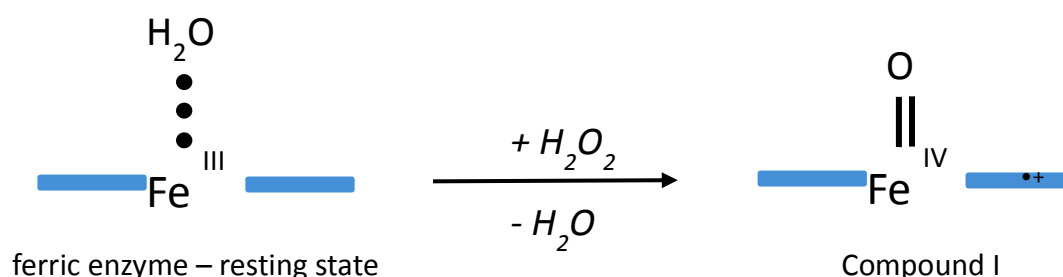


Figure 3-5: Reaction scheme of Compound I formation

One of the stopped-flow apparatus syringes (Braun) was filled with 2 μ M of the enzyme hsPxd01 Con 4N. The protein was diluted in 200 mM phosphate buffer, pH 7.4. It had to be considered that the concentration in the cell would decrease by half, as reactants were mixed in equal parts. So the final enzyme concentration was 1 μ M in 100 mM phosphate buffer for all measurements. The other syringe was filled with the substrate

(H₂O₂). In this case varying concentrations of hydrogen peroxide (Sigma) were used starting with an equimolar concentration. Therefore the final H₂O₂ concentrations were chosen to be: 1 μ M, 2 μ M, 50 μ M and 1 mM.

First measurements were performed by using the photodiode array detector and the conventional mixing mode at the *Stopped flow spectrometer* (Applied Photophysics Sx.18mV-R). A baseline blank was acquired by filling the system with dH₂O. Then spectra in the range of 200 – 750 nm were recorded. The measurements were conducted for 10 seconds. In this time 100 spectra were recorded in a logarithmic scale.

Then measurements using the photomultiplier and therefore recording only one particular wavelength were performed, using the *PiStar CD stopped-flow spectrometer* (Applied Photophysics). Reaction with hydrogen peroxide was followed at 412 nm. The final concentrations in the measuring cell were 500 nM hsPxd01 Con 4N, 100 mM phosphate buffer pH 7.4 and hydrogen peroxide concentrations varied from 2.5 μ M to 10 μ M.

3.4.9.2 Cyanide binding:

Cyanide binds very tightly to the active site of heme proteins in the ferric state. This reaction can be monitored as a spectral shift of the Soret maximum to a higher wavelength.

The stopped-flow syringe was again filled with 2 μ M enzyme solution (hsPxd01 Con 4N) to achieve a final concentration of 1 μ M. A cyanide solution was prepared and final concentrations of 1 μ M, 5 μ M, 10 μ M and 1 mM were measured. As for hydrogen peroxide, spectra were recorded first followed by single wavelength measurements (1 μ M hsPxd01 Con 4N in 100 mM phosphate buffer pH 7.4 and 10 μ M - 100 μ M cyanide) at 434 nm to determine the pre-steady state kinetics of cyanide binding.

3.5 Structural investigations

3.5.1 Size-exclusion chromatography

Size-exclusion chromatography (SEC) can be used as a final polishing step after purification. For example for X-ray crystallography (3.5.2) it is important that the protein is monodisperse (only one species present). SEC was therefore used to separate the monomeric molecules from dimeric or polymeric fractions according to size and shape. The sample passes the column packed with porous medium and smaller parts penetrate the pores where they are retained for different periods of time. Bigger analytes are retained much shorter and thus elute earlier.

The affinity purified constructs were applied to the SEC column before using them for structural investigations. Therefore a SEC *HiLoadTM 16/600 Superdex 200 prep grade column* (GE Healthcare) was connected to the ÄKTA system. The pressure should never exceed 0.3 MPa and the recommended flow rate is 0.5 – 1 mL/min. First the column was washed with dH₂O before it was equilibrated with the self-prepared SEC buffer containing 20 mM phosphate buffer, pH 7.4 and 500 mM NaCl. As the column volume (CV) is 120 mL at least 240 mL (2 CV) of buffer should be applied. The sample was adjusted to 500 mM NaCl and after a centrifugation step the sample was applied to the column using a sample loop appropriate to the sample volume (GE Healthcare). The volume shouldn't exceed 5 mL as this is the maximal volume to be applied to the column without reducing the resolution of the column. After sample application more SEC buffer was applied and 0.5 mL fractions were collected. The elution of proteins varying in size was monitored at 280 and 410 nm (elution profile). Spectra of the respective fractions were recorded and also an SDS-PAGE was performed. After completed elution the SEC column was washed with dH₂O and stored in ethanol (20%).

The primary sequence of a protein can be predicted relatively easy, for example from the nucleotide sequence of the corresponding gene or directly from the protein by mass spectrometry. The determination of the overall shape or furthermore the secondary, tertiary or quaternary structure of proteins at high resolution, however, can be rather labour-intensive.

3.5.2 X-ray crystallography

One of the most commonly used methods to obtain the secondary and tertiary structure of a protein is X-ray crystallography. This method relies on the fact that X-rays are scattered from the electrons in the molecules. Crystallography therefore provides information about the positions of non-hydrogen atoms. The most important prerequisite for successful crystallography is a well-ordered protein crystal which is able to diffract X-rays strongly enough. Although it is necessary to work with pure and homogenous monodisperse protein, the growth of crystals most notably depends on factors like pH, nature of solvent and precipitant, ligands or presence of ions. Therefore a number of commercially available screening kits exists which make it easy to test a lot of different conditions simultaneously (Wlodawer *et al.*, 2008).

There are no general rules for crystallization conditions. The structure of hsPxd01 is unknown and the protein was never crystallized. Therefore a multitude of different conditions was screened to find an environment where the enzyme will form crystals. As a pre-requisite the protein must not aggregate. Therefore all samples were additionally purified by size exclusion chromatography. Further sample preparation was carried out at the Max F. Perutz Laboratories (University of Vienna).

As there is no information available about crystallizing human peroxidase 1, different pre-tests were carried out to learn more about the stability of this enzyme. In a first *thermal shift assay* – also called Thermofluor Assay – the optimal concentration of the fluorescent dye SYPRO Orange (Thermo Fisher) had to be determined. With increasing temperature and denaturation of a protein more and more hydrophobic regions of the protein core become exposed. SYPRO Orange can interact with these patches leading to an increase of fluorescence.

Protein, 300x SYPRO Orange, dH₂O and 1x PBS buffer were mixed in a 96 well plate. Enzyme concentration was kept constant at 0.5 mg/mL while the concentration of the protein dye was increased from 12x to 60x.

The plates were placed in the iCycler iQ5 (BioRad) where the temperature was increased to 95 °C by 3 °C per minute.

With the determined optimal concentration of SYPRO Orange a buffer screen could be performed. The protein/dye mixture was therefore mixed with a variety of buffers with different pH in a 96 well plate and again screened to find the highest T_m values.

Finally, an additive screen was performed, which is again screening for increased T_m values due to the addition of a variety of compounds.

From these results (optimal buffer conditions and optimal additives) an optimal screen was created (CCS Screen).

Crystals were attempted to be grown by the sitting drop vapour diffusion technique. MRC-3-well-plates from SWISSCI were used. These plates comprise 96 wells with three drop-chambers respectively. Each of these three drops contained a different ratio of protein to reservoir solution (200/150nL, 200/200nL, 200/250nL). The samples were administered to the plates using a Nanoliter pipetting robot (Phoenix). Protein concentration was set to be 8 mg/mL or 10 mg/mL depending on the screen.

Many different screens were used. Additionally to the CCS screen derived from pre-experiments, several commercially available screens were employed. Furthermore, a screen based on the conditions for the crystallization of LPO (LPO *opt.*) was set up.

All plates were sealed properly and stored at 22 °C. Progress was monitored either by using the crystal imaging system (Rigaku) or manually on site by microscope.

Screens:

➤ **Low ionic strength (LIS) screen** (Hampton Research):

This screen covers a pH range of 2 to 12 (2, 3, 3.5, 4, 4.5, 5, 5.5, 6.0, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 11, 12) at low ionic strength versus varying PEG 3350 concentrations (4 to 24 % w/v). PEG is therefore used as precipitation agent.

https://hamptonresearch.com/product_detail.aspx?cid=1&sid=25&pid=6

➤ **Morpheus** (Molecular Dimensions):

This 96-condition initial screen combines different approaches, such as chemically compatible mixes of potential ligands, varying buffer systems and mixtures of precipitants. Molecules which were frequently perceived to co-crystallize with proteins (in Protein Data Bank, PDB) were selected and mixed depending on their chemical behaviour and structure. Suitable buffers and precipitants, derived from PDB, were added (Gorrec, 2009).

➤ **JCSG-plus** (Molecular Dimensions):

The JCSG+ screen comprises an optimized sparse matrix with different features such as reduced redundancy, screening of classic PEG and salt conditions, a wide pH range (4 – 10), usage of different organic and polyalcohol conditions as well as neutralised organic acids like formate, acetate or citrate.

<http://www.moleculardimensions.com/shopexd.asp?id=2541>

➤ **Index** (Hampton Research):

This grid screen samples a series of specially formulated reagents including traditional salts versus pH, neutralized organic acids, high salt concentration mixed with low polymer concentration (or high polymer concentration mixed with low salt concentrations), as well as PEG and salt versus pH.

https://hamptonresearch.com/product_detail.aspx?cid=1&sid=24&pid=5

➤ **Crystal Screen** (Hampton Research):

Crystal Screen, which is also designed as sparse matrix, also comprises 96 mixtures of varying pH, salts, polymers and organics.

https://www.hamptonresearch.com/product_detail.aspx?cid=1&sid=17&pid=1

➤ **Salt Rx** (Hampton Research):

For this screen a matrix of varying salt concentration versus pH was established, whereby salt functions as the only precipitant. A pH range of 4 – 9 is covered.

https://hamptonresearch.com/product_detail.aspx?cid=1&sid=32&pid=12

➤ **Wizard Classic crystallization screen** (Rigaku):

The Wizard Classic preparations also include a broad range of crystallants, buffers, and salts as well as a pH range of pH 4.5 to 10.5

<http://www.rigakureagents.com/p-1-wizard-classic-crystallization-screen-series.aspx>

➤ **PACT-premier™** (Molecular Dimensions):

This systematic screen features three different types of screen in one plate: pH/PEG, cation/PEG and anion/PEG. The goal is to test the effect of pH within a PEG/ion environment whereas PEG is used as precipitant.

<http://www.moleculardimensions.com/shopdisplayproducts.asp?id=342>

➤ **LPO Opt:**

This screen was individually designed for the crystallization of the related chordata-enzyme lactoperoxidase, which could be successfully crystalized under similar conditions.

➤ **CCS:**

As described previously, different pre-experiments were made to determine optimal buffer and pH ranges (columns 1 – 12) as well as optimal additives (rows A – H) for human peroxidase 1 variants. Figure 3-6 shows the final screen (CCS) derived from these thermofluor experiments.

		1	2	3	4	5	6	7	8	9	10	11	12
		citric acid 0.1 M pH 5.5	citric acid 0.1 M pH 6.0	MES 0.1 M pH 6.5	HEPES 0.1 M pH 7.0	Tris 0.1 M pH 7.5	Tris 0.1 M pH 8.0	citric acid 0.1 M pH 5.5	citric acid 0.1 M pH 6.0	MES 0.1 M pH 6.5	HEPES 0.1 M pH 7.0	Tris 0.1 M pH 7.5	Tris 0.1 M pH 8.0
Sodium malonate, pH 7.0	A	AS 2 M glycerol 10 w/v						PEG 4k 15 w/v isopropanol 10 v/v					
Betaine hydrochloride	B												
CaCl ₂ 2H ₂ O	C												
control	D												
Sodium malonate, pH 7.0	E	AS 0.1 M PEG 8k 15 w/v						PEG 20k 10 w/v MPEG 2k 12 w/v					
Betaine hydrochloride	F												
CaCl ₂ 2H ₂ O	G												
control	H												

Figure 3-6: Optimized crystallization screen for human peroxidasin 1 constructs

3.5.3 Small-angle X-ray scattering (SAXS)

Small angle X-ray scattering can be used to elucidate the overall shape of a protein in solution. The sample is therefore exposed to X-rays of a distinct wavelength. Those scatter elastically and thereby produce a spatially averaged intensity distribution which can be transformed into a function of spatial frequency (Baldon *et al.*, 2015).

These measurements were carried out at a synchrotron in Berkeley, USA. Samples were sent on ice at a concentration of 1 mg/mL, diluted in 1x PBS buffer. A portion of the buffer was sent there as well to determine the background signal.

4 Results and discussion

4.1 Cloning

4.1.1 Plasmid preparation

The plasmid for cloning was prepared by removing the insert of an established construct by agarose gel electrophoresis after digestion with the restriction enzymes *NheI* and *NotI*.

As shown in Figure 4-1, two different DNA markers (2 log ladder and 1 kb ladder) were applied on the left and on the right side of the sample for determining the size of the bands. Two pieces emerged from this digestion. The bigger one – which thus runs higher – was the vector with a theoretical size of 4400 bp. The smaller piece was the insert, which in this case was Con 5N, and had a size of about 2100 bp. The bigger fragment was excised from the agarose gel and the DNA was purified.

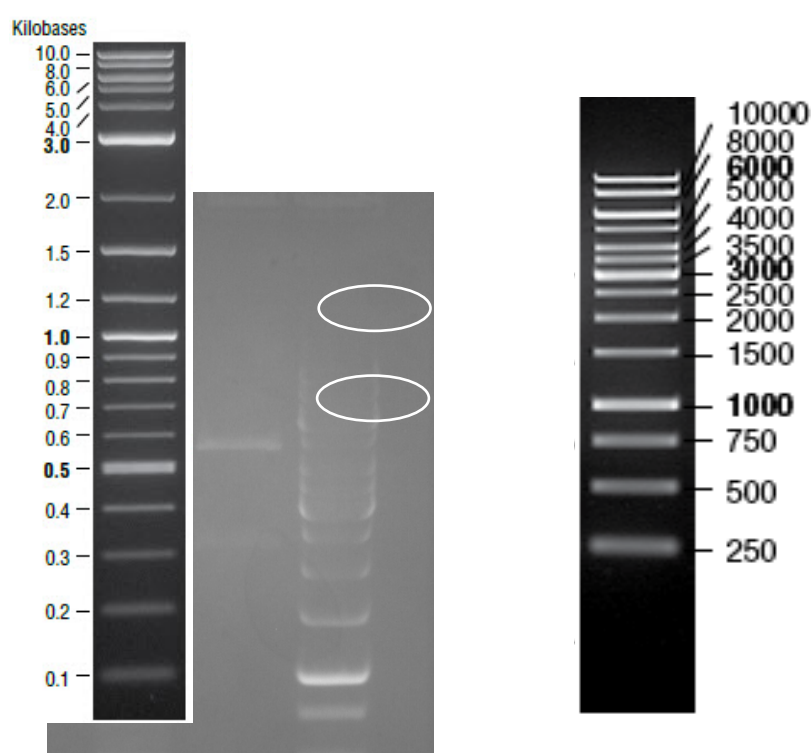


Figure 4-1: Plasmid preparation for subsequent cloning with Gibson Assembly: 2 markers were used: on the left side: 2 log ladder, on the right side 1 kb ladder.

4.1.2 Insert preparation

In order to obtain the distinct inserts for cloning into to pTT5+ vector a miniprep of a former full length construct was used as a template for PCR amplification. Figure 4-2 shows the amplified inserts of constructs Con 1N, 3N and 5N_VWC resolved by agarose gel electrophoresis. The respective bands were excised from the agarose gel and purified.

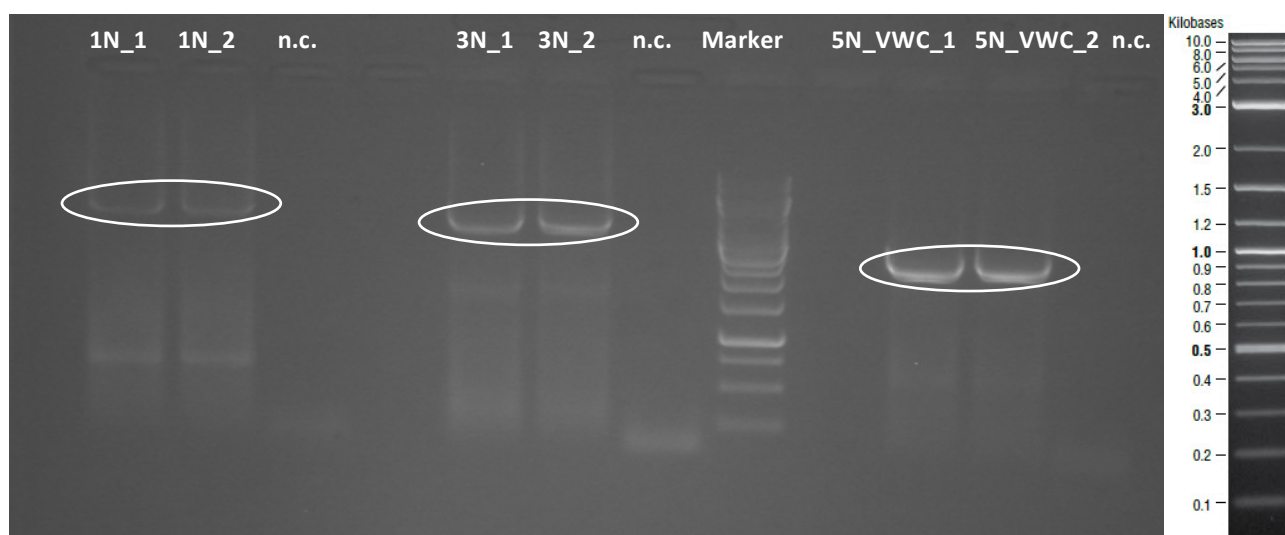


Figure 4-2: PCR product of Con 1N, Con 3N and Con 5N_VWC. n.c.: negative control, Marker: 2 log ladder

4.1.3 Screening

Different assays of screening were used to determine the effectiveness of cloning and to validate the resulting constructs (see 3.1.4).

Figure 4-3 exemplarily shows the result of a restriction enzyme screening. The two restriction enzymes *EcoRI* and *SspI* were used to cut Con 1N. These single cutting enzymes are responsible for the formation of two fragments with theoretical sizes of 5790 and 2997 bp. In this case, clones 8 and 9 appeared to fit best to these conditions and were therefore used for further steps.

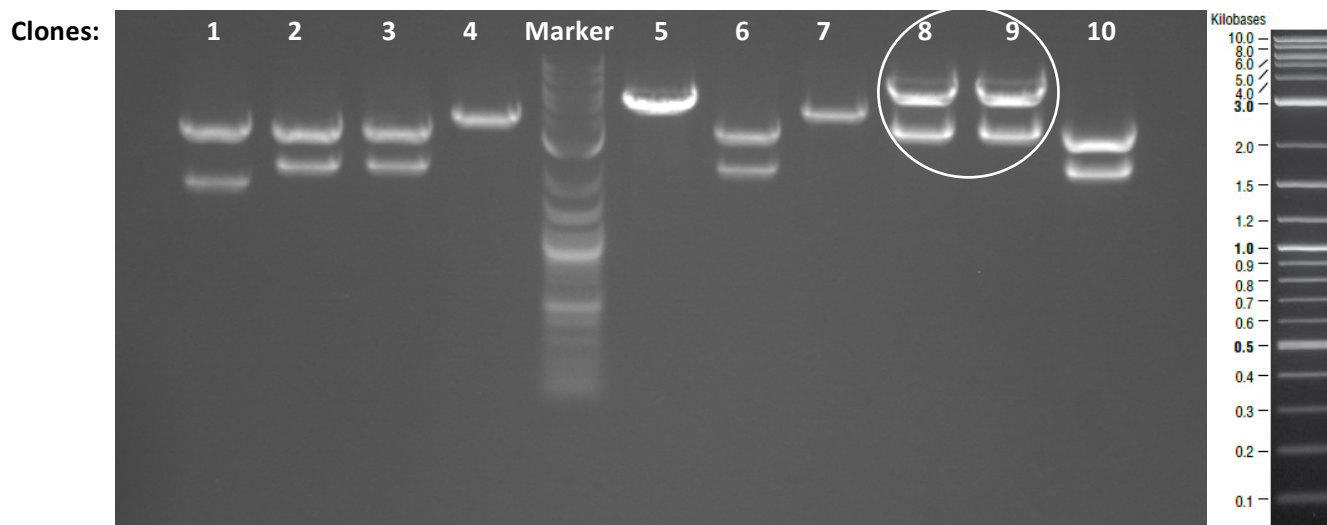


Figure 4-3: Restriction enzyme digest of mini prep DNA of Con 1N clones. Marker: 2 log ladder; clones 8 and 9 were used for sequencing and further analysis.

4.2 Optimization of protein purification

4.2.1 Preparation of supernatant

After thawing the harvested supernatant of cell culture expression (see 3.2.7), it was usually further processed immediately. However, it was observed that an incubation time before diafiltration enhances the correct covalent heme integration into Con 4N significantly. Considerations hereof are that the formation of the covalent linkages, which attach the heme to the active site, might take more time to form completely by traces of hydrogen peroxide which is usually present in all solutions under aerobic conditions. So the thawed sample was stirred at 4 °C for 24-48 hours before the sample was concentrated and buffer exchange took place. This effect could so far only be observed with the variant Con 4N and it was reproducible.

4.2.2 IMAC

As described in section 3.3.2 the diafiltrated sample was purified using an FPLC system (ÄKTA). Figure 4-4 shows a representative profile of an optimized purification of Con 4N. Collection of 2 mL fractions was started with the first gradient. Therefore the concentration of buffer B was initially raised from 0 % to 8 % in 10 minutes. It was held at this ratio until the first peak was eluted completely. Fractions eluting at this relatively low imidazole concentration mainly contained non-specifically bound proteins with a strong reddish-brown colour.

As soon as the absorptions at 280 nm and 410 nm were again at base line level, the second gradient was started. Imidazole concentration was raised to 70 % in a time range of 50 minutes. As imidazole is a stronger binding partner to histidine than nickel is, even strongly bound proteins are eluted from the column. Fractions collected during the elution were coloured greenish due to the heme content of Con 4N.

To determine which fractions contain the highest content of the desired protein, further examinations of the eluted fractions were necessary. UV-Vis spectra were recorded and SDS-PAGE, ECL and Western blot were performed with the fractions resulting from elution of the second peak (see next chapter).

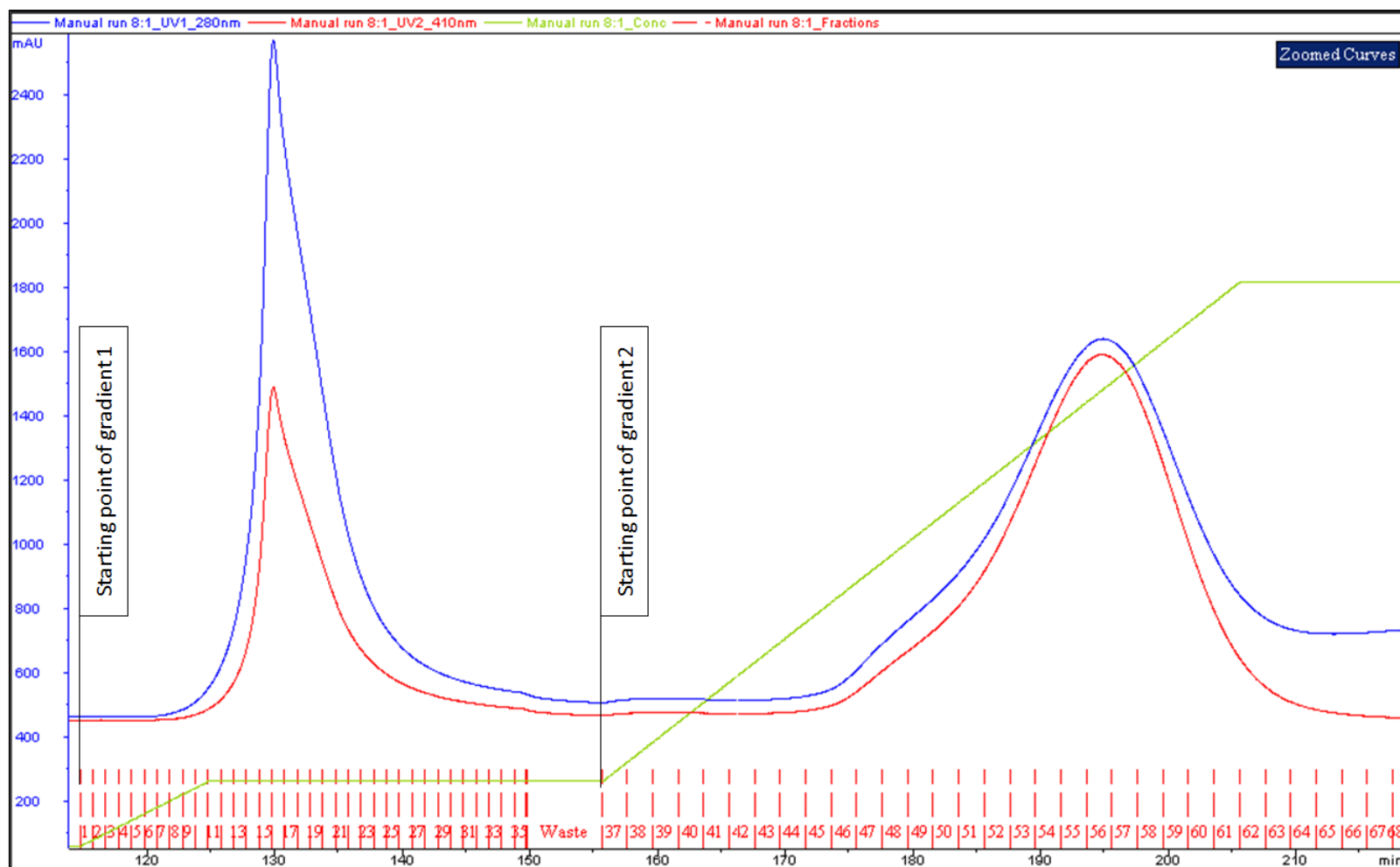


Figure 4-4: HisTrap purification profile of hsPxd01 Con 4N as derived from the ÄKTA system: Two HisTrap FF crude 5 mL columns were coupled, 150 mL sample was applied using a flow rate of 1.5 mL/min. The gradients are formed by buffer A(pH 7.4) and buffer B (pH 7.4), composition see chapter 3.3.2. **blue line:** absorption at 280 nm, **red line:** absorption at 410 nm, **green line:** concentration of buffer B; **bottom:** collected fractions

4.2.3 SDS-PAGE, ECL and Western blot

As described before, the best fractions had to be chosen after elution from the HisTrap column (see 4.2.2). Characteristics of an UV-Vis spectrum are described in chapter 4.3.1. Purity and heme occupancy could additionally be determined by SDS-PAGE, ECL and Western blot.

Fractions were chosen according to the ÄKTA profile (Figure 4-4) and applied to a gel which was further used for ECL measurements and blotting. Figure 4-5 exemplifies the procedure of one purification of Con 4N.

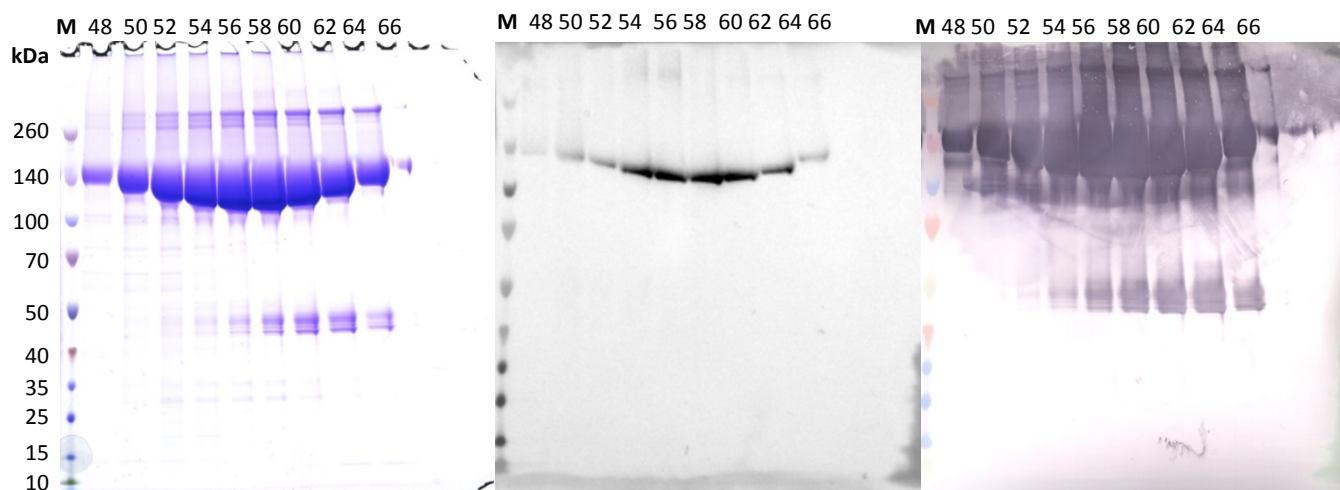


Figure 4-5: SDS-PAGE, ECL and Western blot of fractions obtained from an IMAC purification of Construct 4N: numbers of the lanes refer to fractions seen in the ÄKTA profile (4.2.2). Antibodies used for Western blot: Primary antibody: Penta His Antibody, Mouse monoclonal IgG₁; Secondary antibody: Anti-Mouse IgG. Protein marker VI (M).

Construct 4N doesn't form a trimer and has a theoretical monomeric molar mass of 121 kDa. Due to the high load of protein, bands in Figure 4-5 are very broad. On the SDS-PAGE some degradation products are visible in the later fractions.

ECL reveals covalently bound heme in most of the fractions. There should be two covalent bonds at the active site. So it can be determined whether there is bound heme or not, but if there is one or two bonds present cannot be distinguished.

An Anti-His-tag Western blot also specifically confirmed the presence of peroxidase Con 4N. As micro molar concentrations of protein are required for ECL and the same blot was subsequently used for Western Blotting, the blot is very overloaded.

The distinct fractions with the comparatively best results for these methods, concerning high purity and heme content, were pooled.

The different variants achieved different results from purification concerning amount and purity. Table 4-1 provides an overview over the theoretical purity numbers (describing 100 % heme occupancy) as well as the average RZs achieved by purification (determination will be described in detail in the next section). The dataset was not available for all cloned constructs as not all of them could be expressed and purified in a sufficient amount so far.

Table 4-1: Comparison of the *Reinheitszahl* of 4 different variants: RZ was determined by UV-Vis spectroscopy of purified proteins. The average yield obtained for the different variants is shown in mg/L supernatant.

Variant	Theoretical RZ	Achieved RZ	heme content	mg/L
Con 3N	0.62	0.38	61 %	2.6
Con 4N	0.69	0.40	58 %	19.1
Con 5N	0.76	0.59	78 %	1.9
Con 5N_VWC	0.81	0.53	65 %	2.5

4.3 Protein characterisation

4.3.1 UV-Vis spectroscopy

UV-Vis absorbance spectra were performed either with undiluted fractions resulting from IMAC or for measurements of already purified proteins with the diluted constructs in 100 mM phosphate buffer at pH 7.4. The theoretical RZ-value describing 100 % heme occupancy varies for each construct as it depends on the extinction coefficients at 280 nm and 410 nm. Furthermore the RZ of a distinct construct varied within different expressions and purifications.

Figure 4-6 shows a typical UV-Vis spectrum of the variant Con 4N after purification, concentration and dialysis. The peak at 280 nm is the typical protein peak mainly due to

the absorbance of the aromatic amino acids, tryptophan, tyrosine and phenylalanine. At 412 nm the characteristic Soret Peak of heme *b* is visible. In the far UV-region the values were multiplied by 10 for better visualization. In this region the so-called Q-bands are visible at 510 nm, 547 nm and 590 nm. Further, the charge transfer band can be seen at around 640 nm.

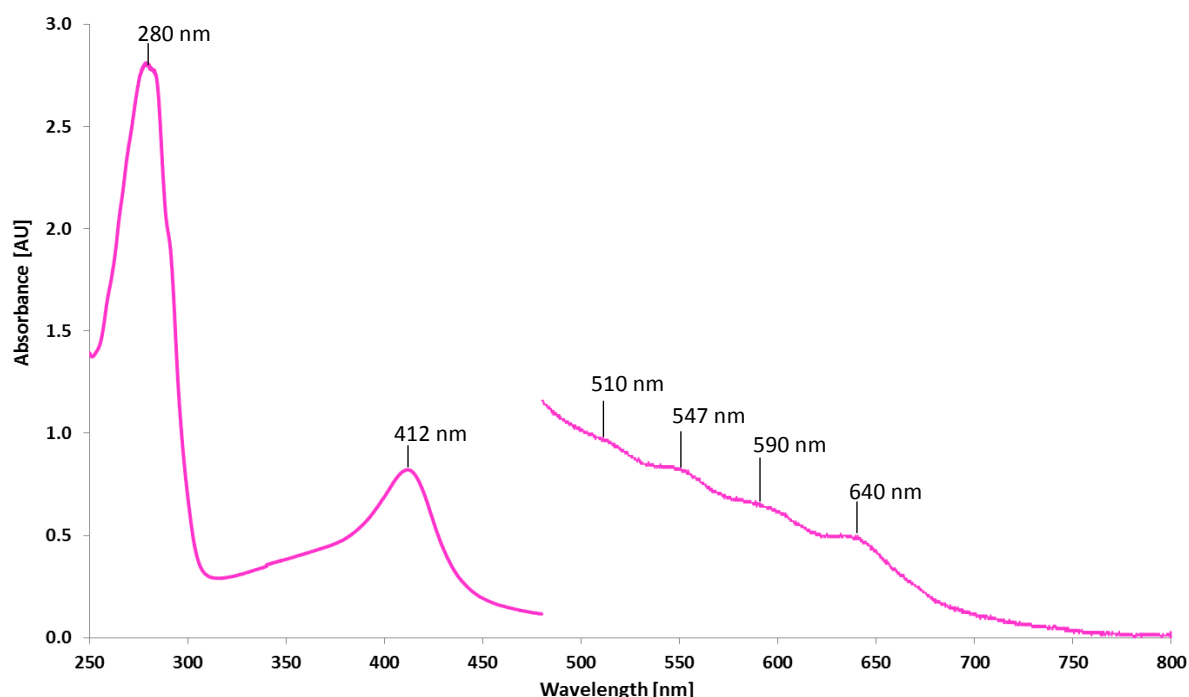


Figure 4-6: UV-Vis spectrum of purified Con 4N in 100 mM phosphate buffer, pH 7.4

Figure 4-7 shows the UV-Vis spectra of the different constructs compared to the chordata peroxidase lactoperoxidase (LPO) which has a very similar active site structure based on amino acid sequence similarity. Further LPO has two covalent heme to protein linkages as would be expected for peroxidasin.

The Soret peak maximum is slightly different for each of the four constructs. Construct 4N shows the highest similarity to the LPO spectrum with a maximum at 412 nm. Construct 3N, 5N and 5N_VWC have their maxima at approximately 410 nm. This could be due to fewer heme to protein linkages, for example only one of the two covalent bonds might actually be established. There are also some differences in the visible region of the spectra as seen on the right side of Figure 4-7. Q-bands and the charge

transfer band of Con 4N are closest to LPO. For the other constructs bands are hardly visible and shifted.

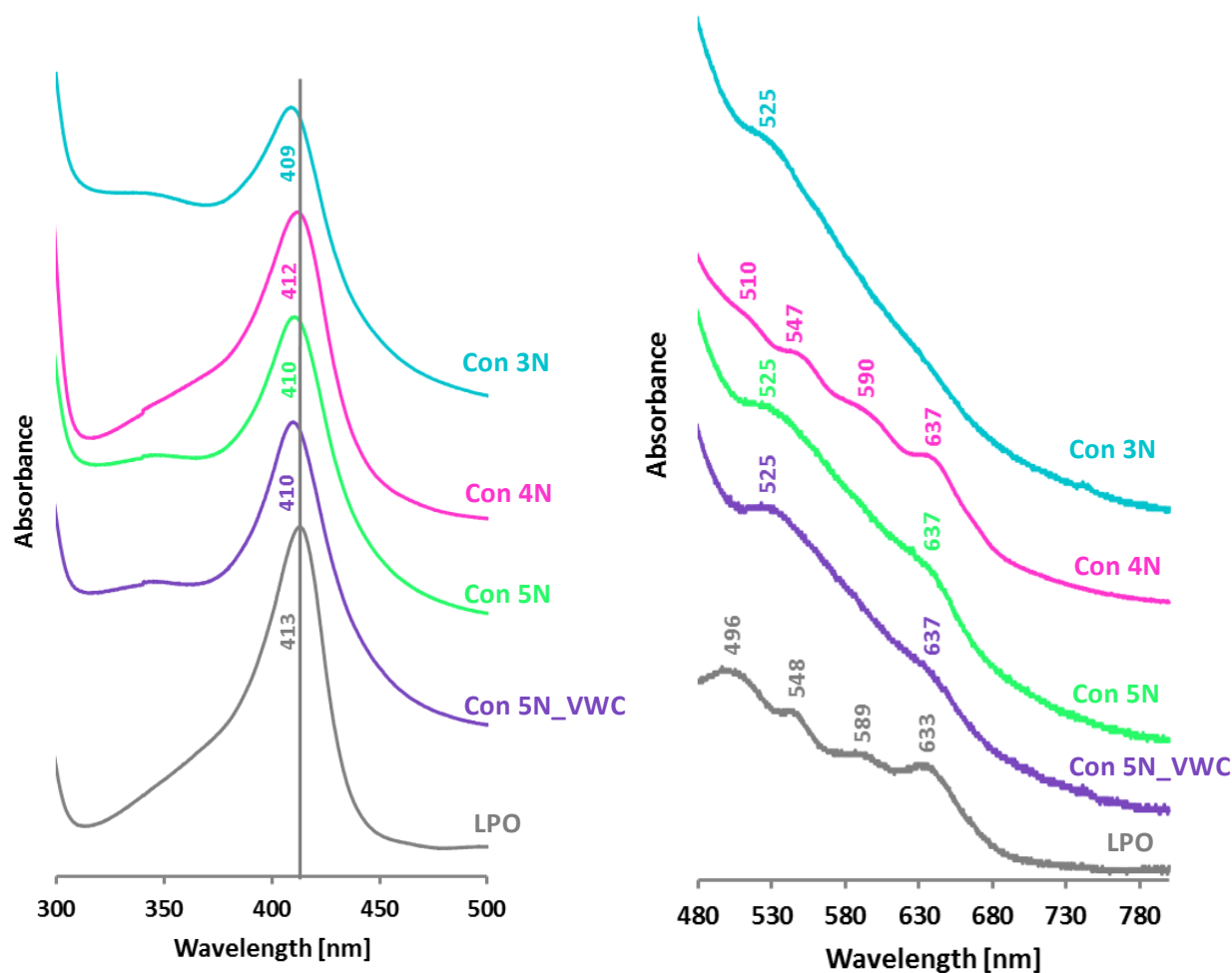


Figure 4-7: Comparison of UV-Vis spectra of Con 3N, 4N, 5N, 5N_VWC and LPO in 100 mM phosphate buffer, pH 7.4: (A) Soret peaks of the enzymes are compared; (B) comparison of the visible region of the constructs

4.3.2 Comparison of constructs

4.3.2.1 SDS-PAGE and ECL

For better visualization of the different sizes of the constructs, an SDS-PAGE of purified constructs was performed. Figure 4-8 shows the comparison of constructs 3N, 4N, 5N

and 5N_VWC under non-reducing as well as reducing conditions. Whereas the first three of them (3N, 4N and 5N) appeared at the height of their monomeric molar masses, Con 5N_VWC formed a trimer which could be reduced.

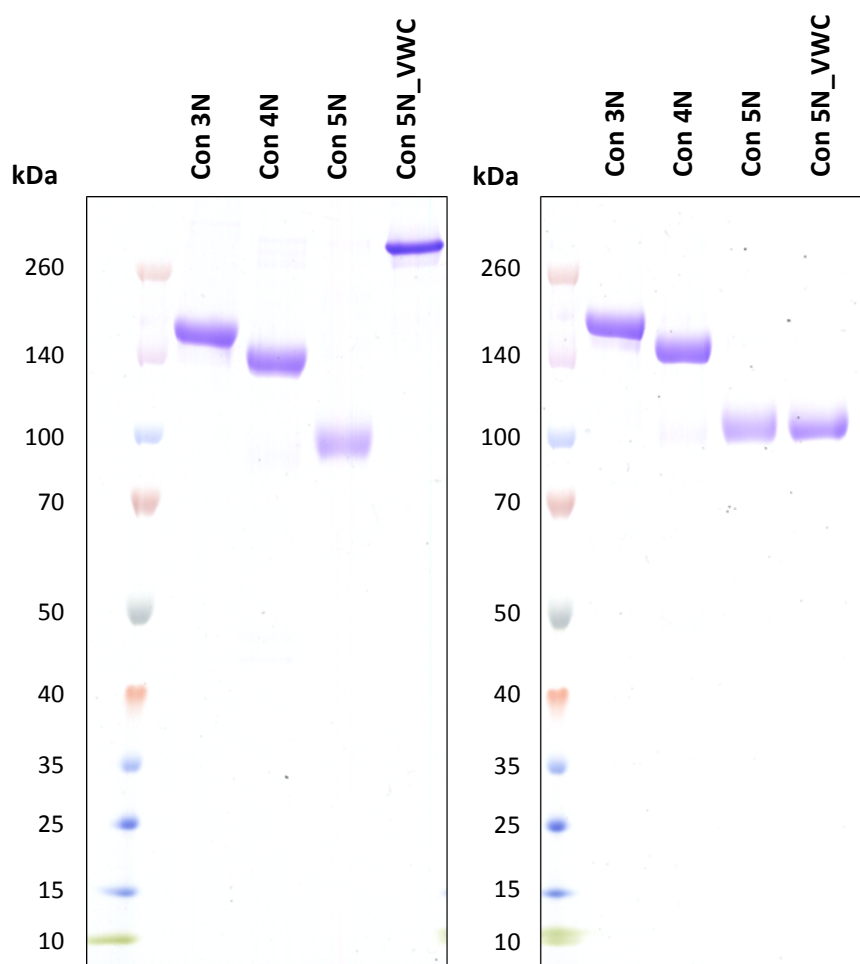


Figure 4-8: SDS-PAGE of Con 3N, 4N, 5N, 5N_VWC: non-reducing (left) and reducing (right) gel, revealing the trimerization of construct Con 5N_VWC

Samples of these four constructs were also used to perform ECL measurements (Figure 4-9). As a result one can see that Con 4N, 5N and 5N_VWC give clear signals at the expected molar mass comparable to the corresponding SDS-PAGE (Figure 4-8). Despite the typical heme-spectrum for construct 3N, as seen before, the blot reveals that most of the heme is not covalently linked.

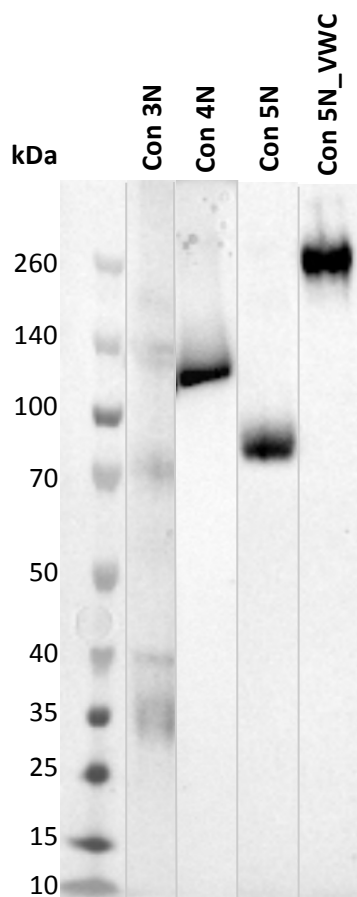


Figure 4-9: ECL measurement of purified Con 3N, 4N, 5N and 5N_VWC

Because of this result and the UV-Vis spectral data, it was assumed that – at the moment – Con 4N is most likely the only variant which contains a heme with two covalent bonds. That's why the majority of the following characterization data was accomplished with Con 4N.

4.3.3 Investigations on the C-terminus

It was already known that human peroxidase 1 forms a homotrimer (Soudi *et al.*, 2015). To further investigate the oligomerization state and the domains involved in it, several different constructs were expressed and compared (Lázár *et al.*, 2015).

SDS-PAGE and Western blot revealed that the constructs Con 3N, Con 4N and Con 5N are monomers as they run at the expected molecular weight of their monomeric sizes. The variants Con 1N as well as Con 5N_VWC form a trimer (see 4.3.2.1).

So it was concluded that the C-terminal region, lacking in Con 3N, Con 4N and Con 5N, as they end directly after the peroxidase domain, must be involved in the oligomerization.

A closer look at the amino acid sequence of the linker region – connecting the POX domain with the VWC domain – reveals that the peroxidase domain is followed by a short stretch of 3 cysteine residues with a subsequent amphipathic helical sequence. Based on this, the two variants Con 5N_cys and Con 5N_helix were produced, containing either just the cysteine residues or the helix in addition (see chapter 3.1).

Small scale expression and subsequent analysis using SDS-PAGE, ECL and Western blotting of non-purified HEK cell supernatant revealed that the VWC domain is not involved in the trimerization as Con 5N_helix already appeared as a single redox sensitive trimer band. In contrast, Con 5N_cys appeared mainly in monomeric form with little dimer and trimer formation.

Under reducing conditions in the presence of DTT, which is able to reduce disulphide bonds, both constructs appear as monomers (see Figure 4-10).

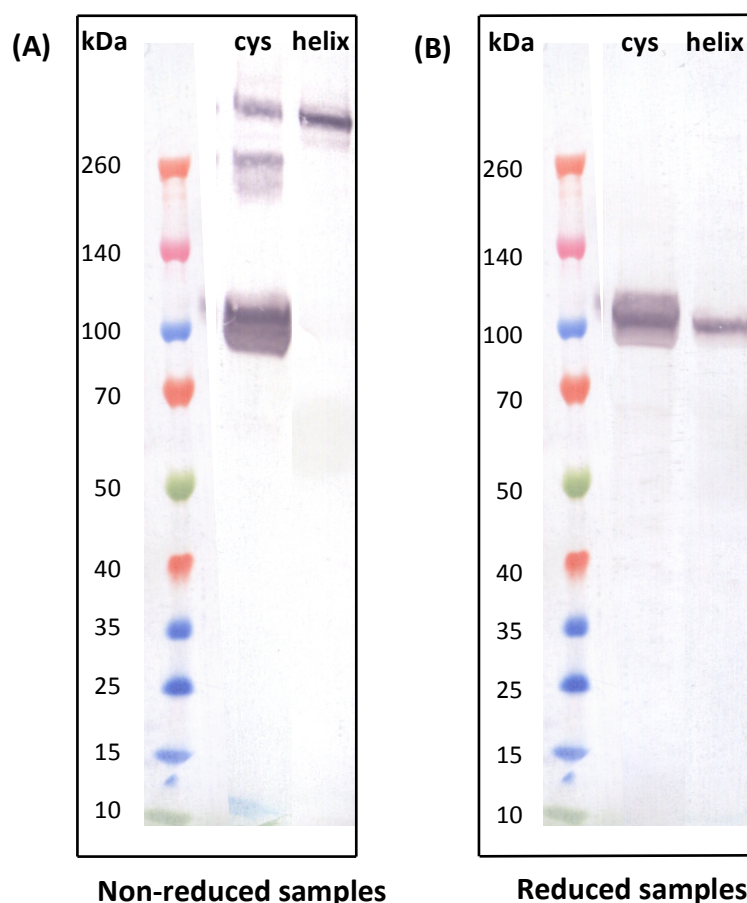


Figure 4-10: Western blot of variants Con 5N_cys ("cys") and Con 5N_helix ("helix") under non reducing conditions (A) and reducing conditions (B) revealing the sites of oligomerization

Based on these results it is assumed that the amphipathic helical region is essential for an optimal steric arrangement of the monomers. It might thereby assist disulphide bridge formation between monomers. So the establishment of this triple helical structure may be a requirement for correct orientation. Apparently the VWC domain is not involved in the oligomerization.

4.3.4 Circular dichroism spectroscopy

Recording of CD spectra in the visible region should give insights into the immediate heme environment. In addition to the four variants (Con 3N, 4N, 5N and 5N_VWC) LPO was again used as a model enzyme, so the spectra of the variants can be compared.

The spectra were normalized concerning the heme content of the variants and are depicted in Figure 4-11.

Con 4N showed the most similar signal when compared to LPO. This confirmed previous results where this variant was shown to have an LPO like heme environment. The other constructs measured, not just showed a highly reduced ellipticity and in general a rather weak signal which is in contrast to LPO and Con 4N, but also their maxima were shifted to higher wavelengths. All these results indicate a different, maybe not correctly orientated heme surrounding of these variants (3N, 5N, 5N_VWC).

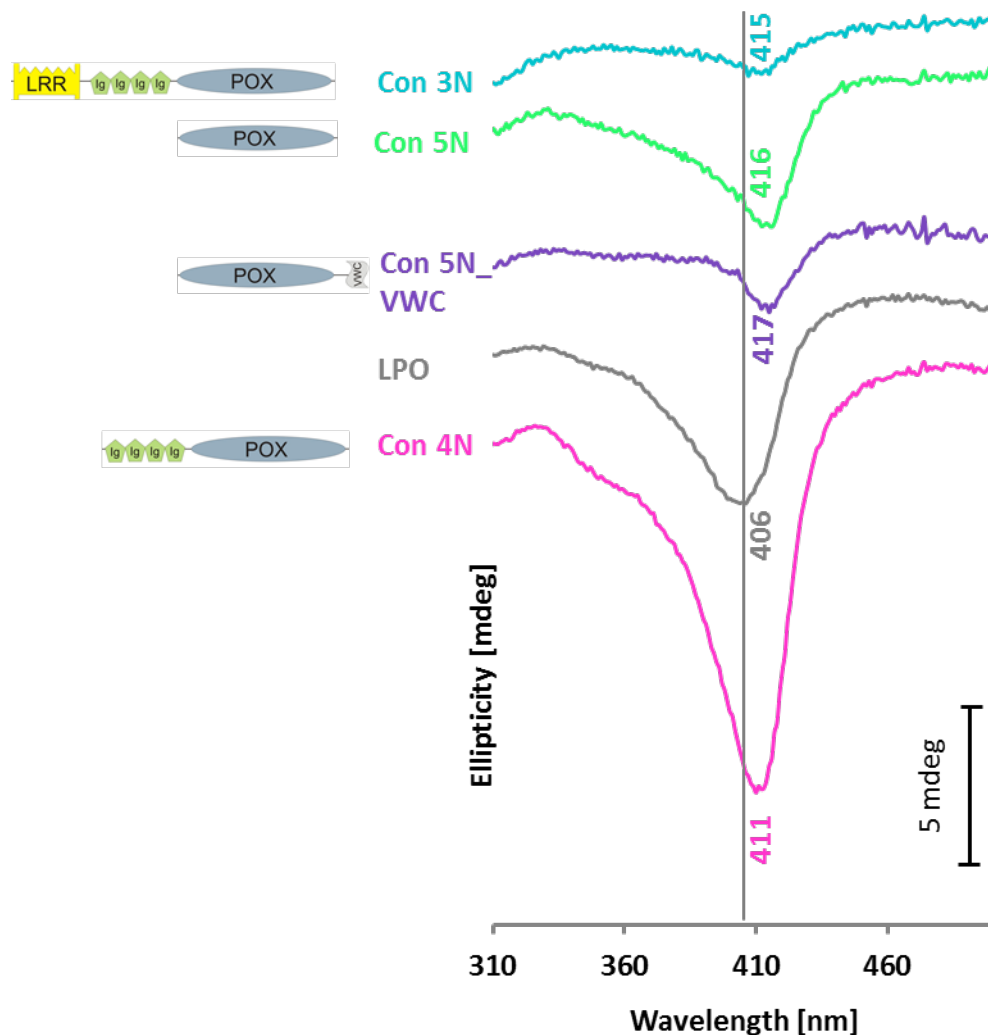


Figure 4-11: CD spectral data of variants Con 3N, 4N, 5N and 5N_VWC compared to LPO; all samples measured in 100 mM phosphate buffer pH 7.4

4.3.5 Spectroelectrochemistry

The presence of covalent heme to protein linkages has an influence on the heme distortion. This further has an impact on the redox properties of the protein. Therefore spectroelectrochemical determination of the standard reduction potential of the Fe(III)/Fe(II) couple was performed.

The spectral changes obtained by the redox titration – whereby the ferro form was turned into the ferri form – were used to determine these standard reduction potentials of the constructs. Redox titration of Con 4N lead to a shift of the Soret maximum from 413 nm (oxidized form in ferric state) to 435 nm (reduced form in ferrous state) which is a typical value for a 5-coordinated high-spin heme.

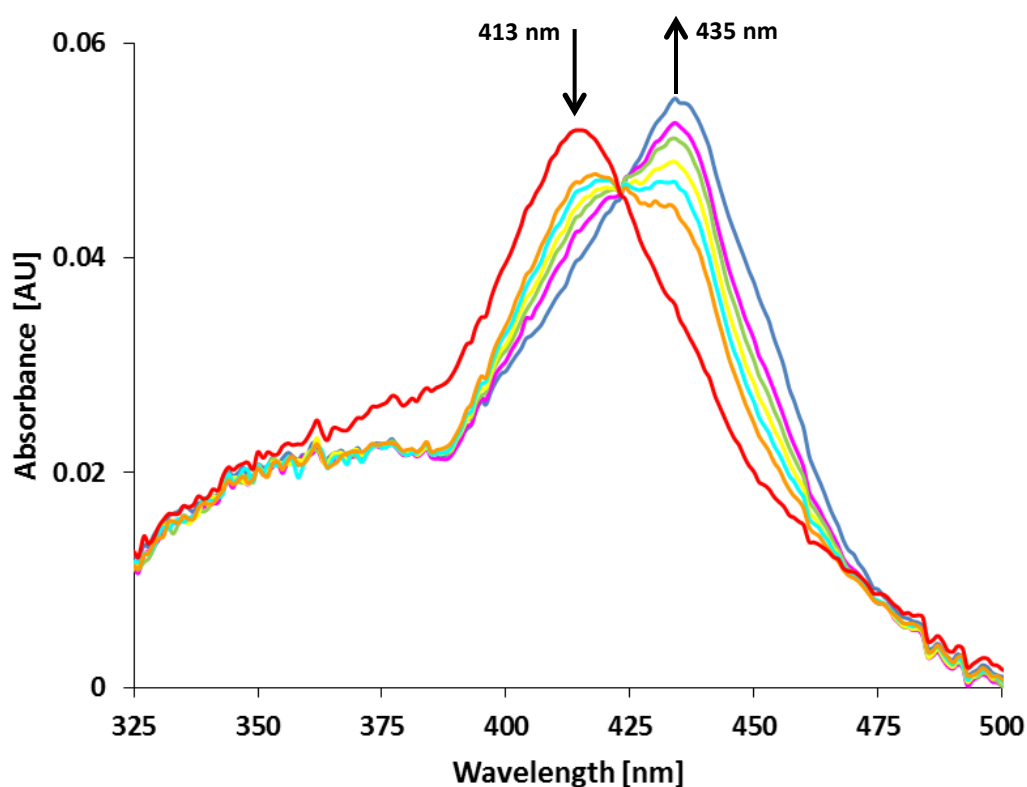


Figure 4-12: Redox titration of Con 4N: a shift in the Soret maxima from 413 nm to 435 nm is visible.

Upon measuring the equilibria $K = \frac{[\text{Ox}]}{[\text{Red}]}$ at distinct applied potentials, a Nernst plot can be obtained. The slope of the linear curves (here 80.99) functions as a quality control. The optimal value would be 58. The more the obtained value varies from this, the more inexact is the result. The present one can be considered as a quite reliable result.

The intercept of -128.38 for Con 4N equals the standard reduction potential in Millivolt.

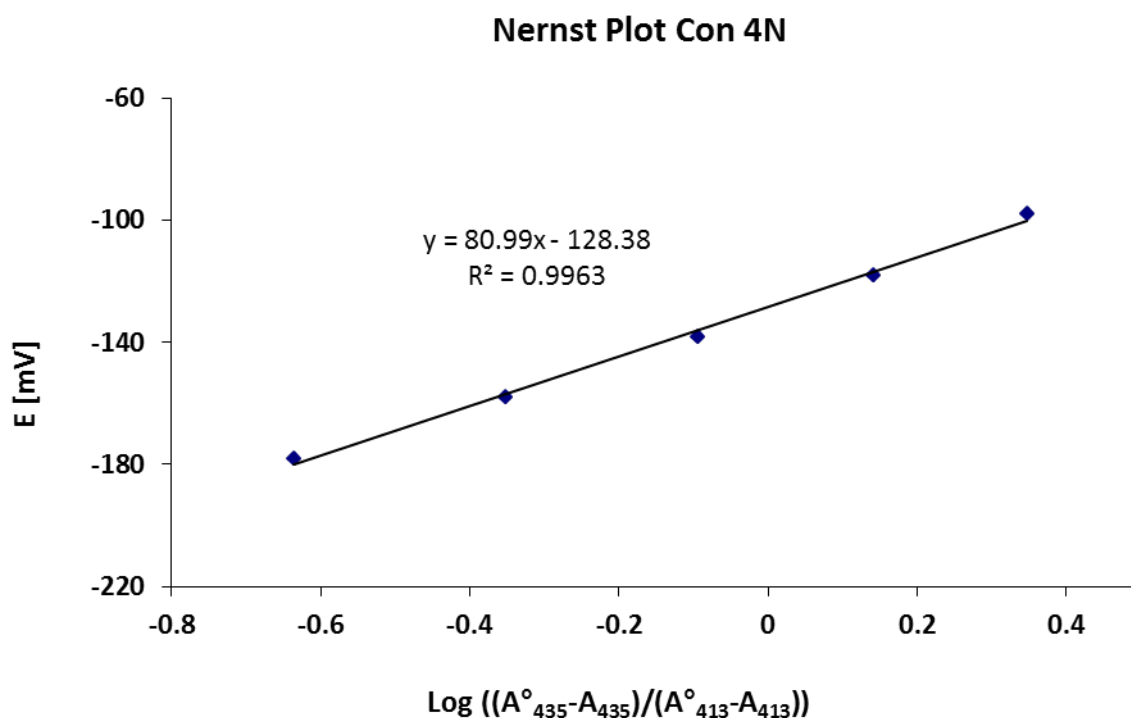


Figure 4-13: Nernst Plot derived from redox titration of Con 4N: points were connected by a trendline to obtain the linear equation.

The resulting values (listed in Table 4-2) were again compared to LPO. Con 4N showed the most positive redox potential of the constructs and is even higher than that of LPO. This would explain the ability of peroxidasin to oxidate also chloride to a small extent in addition to iodide and bromide.

Table 4-2: Standard reduction potential of the redox couple [Fe(III)/Fe(II)] of various peroxidasin constructs compared to LPO.

Construct	E° [mV]
Con 3N	-195 ± 10
Con 4N	-128 ± 10
Con 5N	-201 ± 10
Con 5N_VWC	-198 ± 10
LPO	-176

4.3.6 Activity measurements

4.3.6.1 Peroxidase activity

Peroxidase activity was detected spectrophotometrically using hydrogen peroxide and ABTS or guaiacol as substrates.

ABTS oxidation:

Preliminary results showed that Con 4N was the most active variant. Also it is the variant with the highest expression yield. Therefore Construct 4N was used to determine the optimal pH of the assay. The absorbance measured was converted to the specific activity and plotted against the distinct pH value (Figure 4-14).

The calculation of the specific activity [U/mg] is exemplarily shown for one data point of Con 4N at pH 5:

Specific activity: $A = [\text{Units mg}^{-1}] = [\mu\text{mol ABTS min}^{-1} \text{mg}^{-1} \text{protein}]$

$\epsilon_{\text{ABTS ox 414nm}} = 31100 \text{ M}^{-1}\text{cm}^{-1}$

Absorbance: 0.1149 AU/min

$\text{Abs/min} / \epsilon = 3.7 \cdot 10^{-6} \text{ mol/min} \cdot \text{L} = \mathbf{3.7 \mu\text{mol/min} \cdot \text{L}}$

$3.7 \mu\text{mol/min} \cdot \text{L} * \text{Volume used} = 3.7 \mu\text{mol/min} \cdot \text{L} * 0.0008 \text{ L} = \mathbf{2.96 \cdot 10^{-3} \mu\text{mol/min}}$

$2.96 \cdot 10^{-3} \mu\text{mol/min} / \text{mg used} = 2.96 \cdot 10^{-3} \mu\text{mol/min} / 0.00484 \text{ mg} = \mathbf{0.611 \text{ U/mg}}$

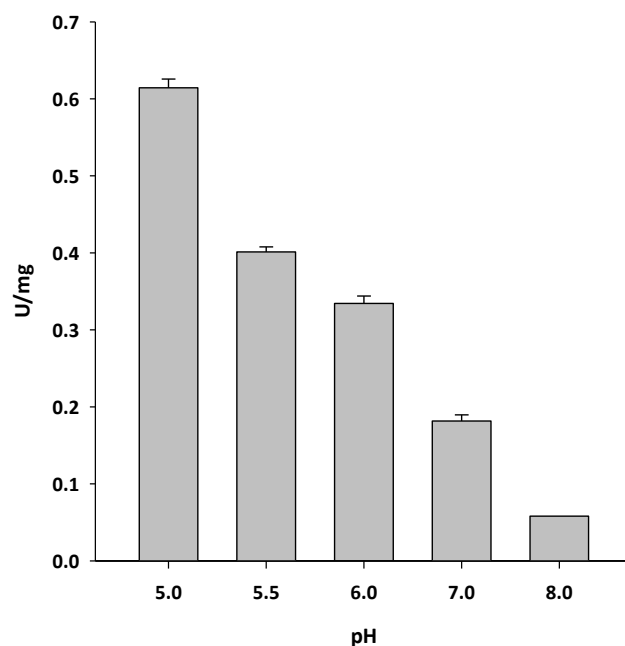


Figure 4-14: Con 4N ABTS assay: pH dependence

As seen in Figure 4-14 the highest ABTS activity was measured at pH 5.

The ABTS activity of Con 3N, 5N and 5N_VWC were subsequently measured under the same conditions at the determined optimal pH. Con 4N showed by far the highest activity (Figure 4-15) although in general the overall specific activity with ABTS was relatively low compared to other peroxidases. Table 4-3 shows the distinct values for the specific activities of the variants Con 3N, 4N, 5N and 5N_VWC compared to each other.

At this point it has to be mentioned that the steady state activity measurements of Con 4N presented here were performed before the purification of Con 4N was even more optimized (inclusion of a 24-48 h incubation step before purification). As mentioned above this incubation step led to an increased activity presumably due to further covalent heme to protein link formations. This means that the specific activity of Con 4N depicted here is lower than it would be with the latest purification protocol. For time reasons these measurements were not repeated during this thesis.

Table 4-3: Specific activity of the constructs with ABTS assay

hsPxd01 variant	Specific activity [U/mg]
Con 3N	0.016
Con 4N	0.61
Con 5N	0.018
Con 5N_VWC	0.06

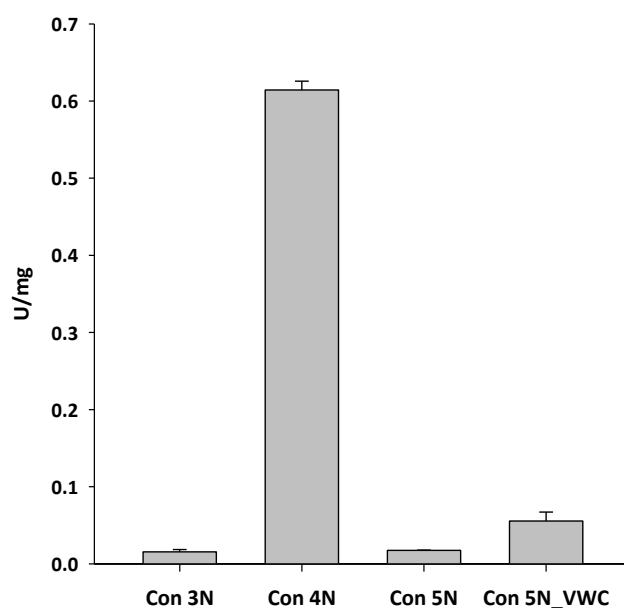


Figure 4-15: Comparison of ABTS assay activity of different constructs at pH 5

Guaiacol oxidation:

The peroxidase activity was also measured with Guaiacol and the specific activity was determined again. The calculation of the specific activity was carried out as shown above for the ABTS assay. Also the mode of the measurements was maintained: first the pH dependency of the overall reaction was determined with Con 4N and then the guaiacol activity of the constructs was determined under the determined optimal pH.

Specific activity: $A = [\text{Units mg}^{-1}] = [\mu\text{mol guaiacol min}^{-1} \text{mg}^{-1} \text{protein}]$

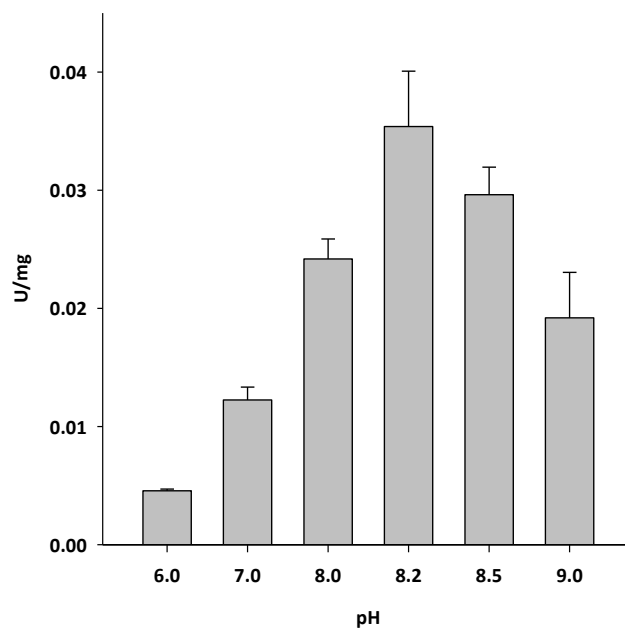


Figure 4-16: Con 4N guaiacol assay: pH dependence

As seen in Figure 4-16 the optimal pH for this assay performed with Con 4N was at pH 8.2. Again, the activity of the other constructs were measured at this pH (see Table 4-4 and Figure 4-17) Con 4N was again the most active construct. However, the specific activity of about 0.025 U/mg is very low compared to other peroxidases and guaiacol seems to be an even worse substrate for peroxidase than ABTS is.

Table 4-4: Specific activity of the constructs for guaiacol assay

hsPxd01 variant	Specific activity [U/mg]
Con 3N	0.006
Con 4N	0.025
Con 5N	0.0008
Con 5N_VWC	0.016

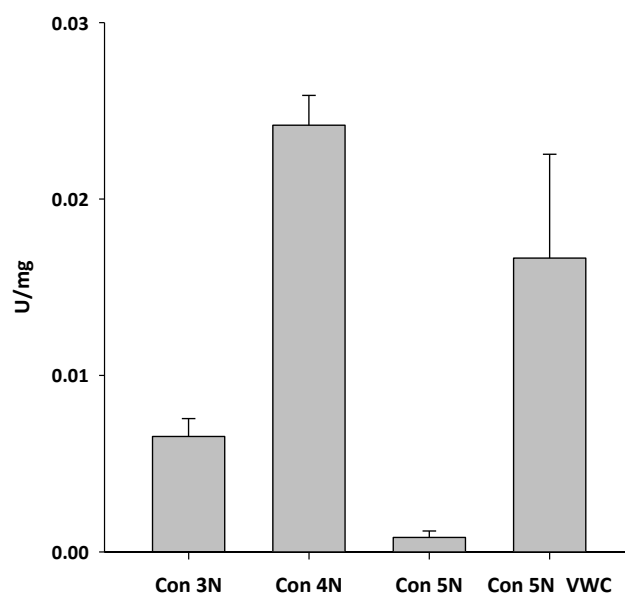


Figure 4-17: Comparison of the constructs at pH 8 using the guaiacol assay.

4.3.6.2 Halogenation activity

Available publications about halogenation activity of hsPxd01 suggest that the activity is relatively low compared to chordata peroxidases although they show a high sequence similarity (Soudi *et al.*, 2015).

Bromination:

The kinetic constants derived from the oxidation of bromide were determined by observing the bromination of NADH forming bromohydrin. Thereby the concentrations of bromide were varied.

First, the optimal pH range for the assay was determined again with Con 4N. The plots were derived from plotting specific activity vs. pH. Calculation of U/mg was performed as described under ABTS assay (see above):

Specific activity: $A = [\text{Units mg}^{-1}] = [\mu\text{mol Br}^- \text{ min}^{-1} \text{ mg}^{-1} \text{ protein}]$

$\epsilon_{\text{Bromohydrin 275 nm}} = 11800 \text{ M}^{-1}\text{cm}^{-1}$

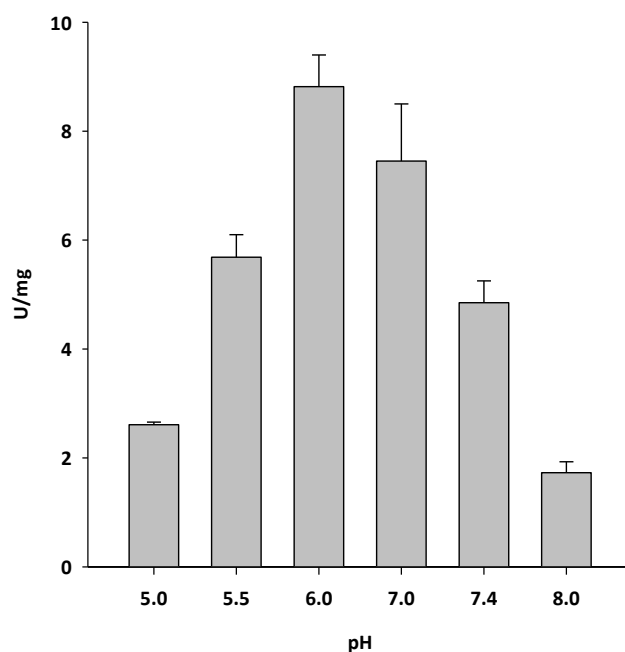


Figure 4-18: Con 4N Bromide/NADH assay pH dependence

As shown in Figure 4-18, pH 6 was the pH with the highest activity. Con 4N also showed almost 10 times increased activity when compared to the peroxidase assays with ABTS and guaiacol. But it also has to be mentioned that these measurements were performed with a newer, even better purification of Con 4N (including the incubation step of 24-48 h) than in the previously described peroxidase assays.

Further, the constructs were again compared to each other under the same conditions at the optimal pH. Results of these measurement are shown in Table 4-5 and in Figure 4-19. Once more, Con 4N is by far the most active construct.

Table 4-5: Bromination activity of the different constructs

hsPxd01 variant	Specific activity [U/mg]
Con 3N	0.03
Con 4N	8.82
Con 5N	0.75
Con 5N_VWC	0.09

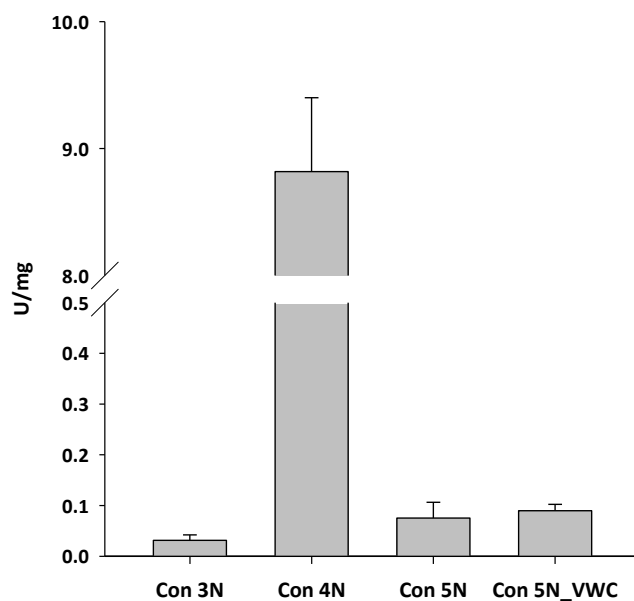


Figure 4-19: Comparison of the bromide/NADH assay for different constructs at pH 6

Because of these results the determination of the kinetic constants was performed with construct 4N only.

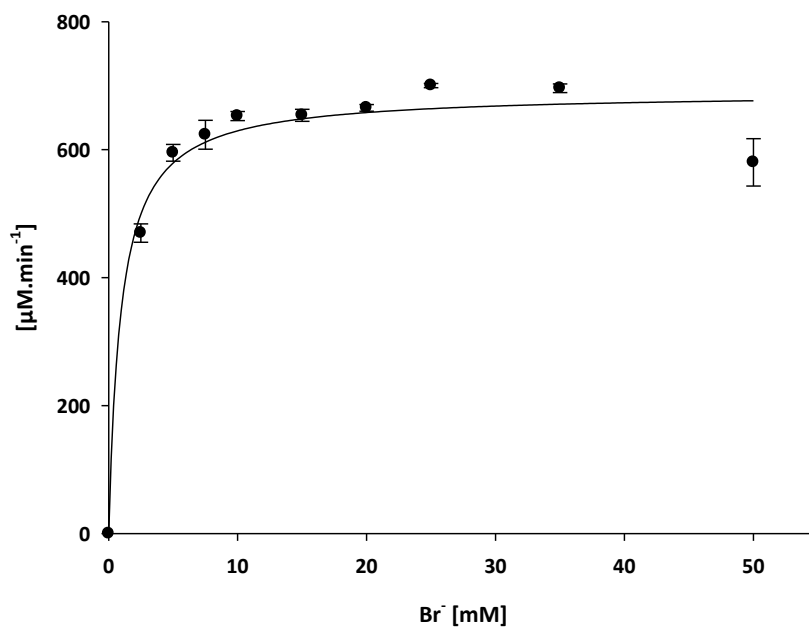


Figure 4-20: Varying bromide concentrations at constant H₂O₂ concentration of 50 μM: the measurements were performed with 100 nM Con 4N purification 16-17 (best one so far) in 20 mM phosphate buffer, pH 7.4 and in the presence of 100 μM NADH.

Figure 4-20 shows the plot derived from measurements with a constant hydrogen peroxide concentration of 50 μM and varying bromide concentration in 20 mM phosphate buffer at pH 7.4. The curve was fitted using *Sigma Plot* and the following values were obtained using the equation:

$$f = y0 + \frac{a * x}{b + x}$$

Resulting in: $a = 689.3561 \pm 24.1030$

$b = 0.9520 \pm 0.3190$

Thereby **a** describes v_{max} , following that $v_{max} = 689.4 \mu\text{M} \cdot \text{min}^{-1}$ which is higher than every value determined for peroxidasin before. This could be due to the optimized purification.

b is the value calculated from the programme for K_M : $K_M = 0.95 \text{ mM}$ with a standard error of $\pm 0.3 \text{ mM}$.

These kinetic constants were also determined using the double reciprocal Lineweaver-Burk plot. Therefore $1/[S]$ was plotted against $1/v$. Figure 4-21 shows the basic principle of such plots. By displaying the linear equation, the constants K_M and v_{max} can be determined easily from the slope and the intercept as depicted below.

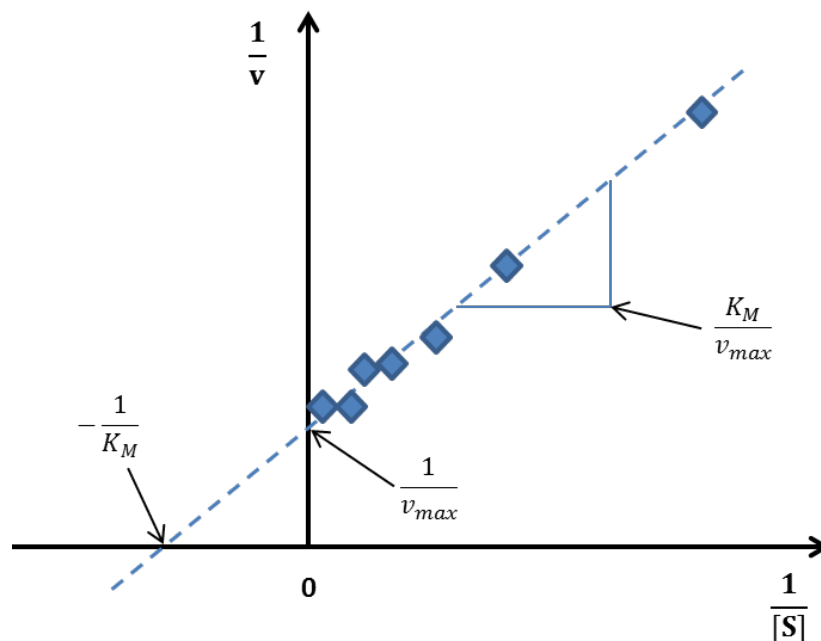


Figure 4-21: Scheme of a Lineweaver-Burk plot: this double reciprocal plot can be used for determination of kinetic constants from the linear equation.

The resulting plot is shown in Figure 4-22. The linear equation derived from the data points is:

$$y = 0.0016 * x + 0.0014$$

Following the scheme of Figure 4-21 the following terms can be determined:

If $x = 0 \rightarrow y = \frac{1}{v_{max}}$

And $y = 0 \rightarrow x = -\frac{1}{K_M}$

With these equations the following values can be calculated:

$$v_{max} = 695.12 \mu\text{M} \cdot \text{min}^{-1}$$

$$K_M = 1.08 \text{ mM}$$

The values determined with the double reciprocal plot therefore fit together very well with those determined by Sigma Plot.

Further, two more constants were calculated:

k_{cat} : the turnover number [sec^{-1}] which incorporates the concentration

$$k_{cat} = \frac{v_{max}}{E_0}$$

With E_0 being the enzyme concentration used (100 nM)

$$k_{cat} = \frac{695.12 \mu\text{M} \cdot \text{min}^{-1}}{0.1 \mu\text{M}} = 6951.15 \text{ min}^{-1}$$

k_{cat}/K_M : a value for measuring the enzyme efficiency

$$\frac{k_{cat}}{K_M} = \frac{6951.15 \text{ min}^{-1}}{1.08 \text{ mM}} = 6.44 * 10^6 \text{ M}^{-1} \cdot \text{min}^{-1}$$

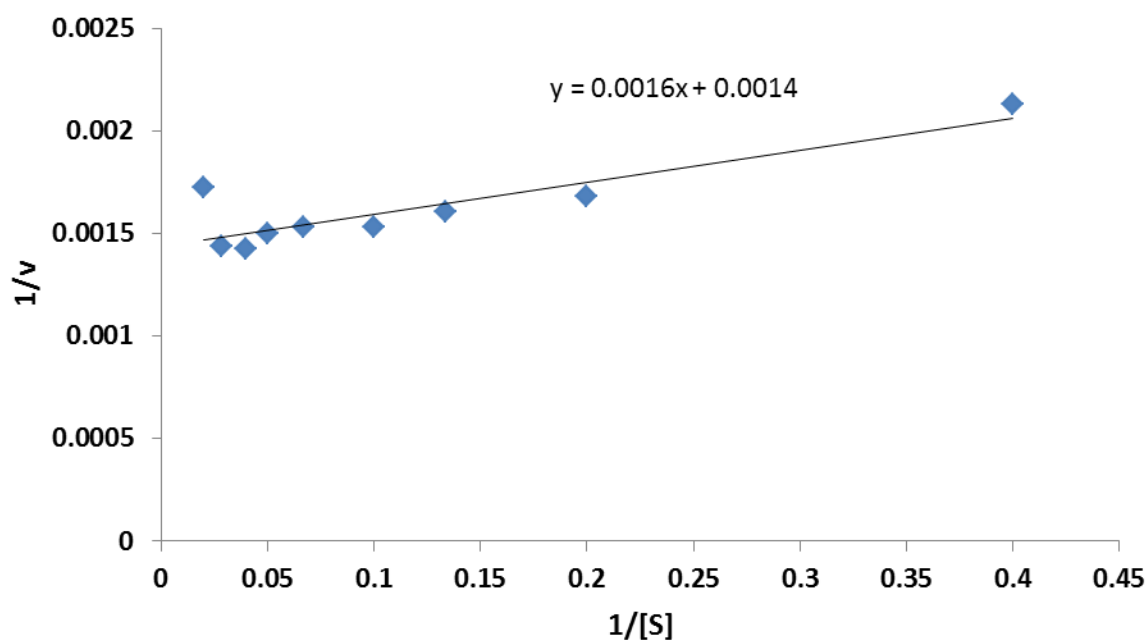


Figure 4-22: Lineweaver-Burk plot for the determination of kinetic constants: the system contains 100 nM Con 4N, 50 μM H_2O_2 , 100 μM NADH and varying concentration of Br^- (0-50 mM).

Table 4-6 summarizes the determined constants.

Table 4-6: Summary of the determined kinetic constants for bromination activity of Con 4N

v_{max}	695.12 $\mu\text{M} \cdot \text{min}^{-1}$
K_M	1.08 mM
k_{cat}	6951.15 min^{-1}
k_{cat}/K_M	$6.44 \cdot 10^6$ $\text{M}^{-1} \cdot \text{min}^{-1}$

Thiocyanate oxidation:

The catalysis of the formation of hypothiocyanate (OSCN^-) could be monitored by employing the 5-thio-2-nitrobenzoic acid (TNB) assay.

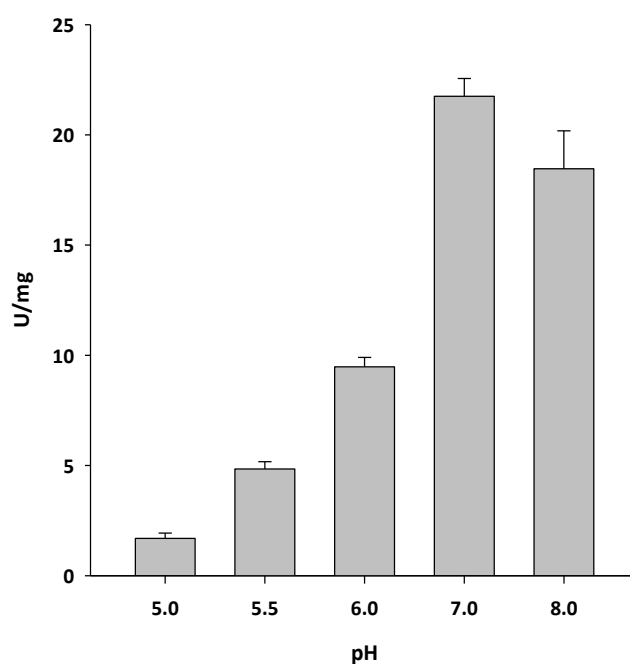


Figure 4-23: Con 4N SCN-/TNB pH dependence: 100 nM enzyme, 200 μ M H₂O₂, 100 mM SCN⁻, 100 μ M TNB and the distinct buffers (100 mM).

As depicted in Figure 4-23 the optimal pH was around pH 7. This was determined with Con 4N. Further, the constructs were again compared to each other. Figure 4-24 shows the result of this comparison and Table 4-7 summarizes the obtained specific activities of the constructs.

Table 4-7: Thiocyanate oxidation activity of the different constructs:
enzymes diluted in 100 mM phosphate buffer, pH 7.

hsPxd01 variant	Specific activity [U/mg]
Con 3N	0.03
Con 4N	21.75
Con 5N	1.10
Con 5N_VWC	0.70

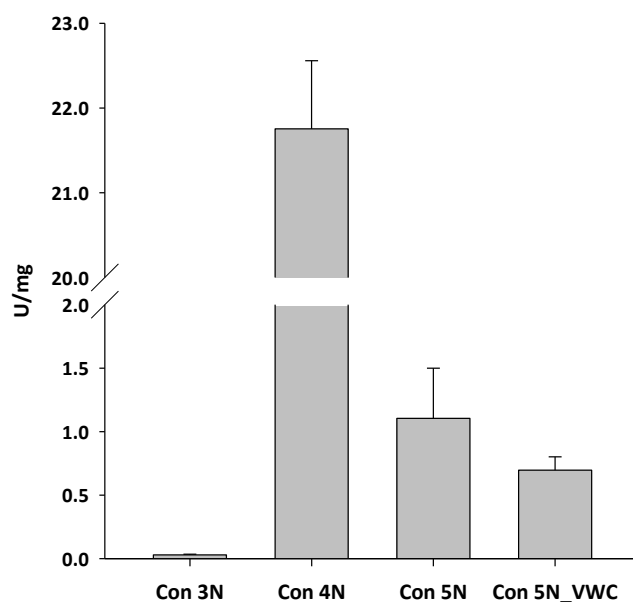


Figure 4-24: OSCN/TNB assay at pH 7: comparison of the constructs.

The oxidation of thiocyanate was considerably faster than the reaction with bromide which was surprising as it was already proven that bromide is the physiological substrate of peroxidasin (McCall *et al.*, 2014). But as it was not possible to achieve reproducible results when trying to vary the thiocyanate concentration (as seen for bromide before) to then determine kinetic constants, further measurements were performed (see 4.3.7).

4.3.7 Hydrogen peroxide consumption

The utilization of hydrogen peroxide by variant Con 4N (purification number 16-17 including the incubation step) was followed polarographically with a hydrogen peroxide sensitive electrode. Therefore the loss of H_2O_2 in the system consisting of Con 4N, H_2O_2 and a (pseudo-)halide was measured.

First of all the reaction was followed when no electron donor was added. About 20 % of hydrogen peroxide were consumed under these conditions as depicted in blue (Figure 4-25 and Figure 4-26).

Subsequently the measurements were performed with typical physiological substrates. Figure 4-25 shows that in the presence of 140 mM sodium chloride the hydrogen peroxide consumption almost doubled. Approximately 70 % of hydrogen peroxide were consumed rapidly when 100 μ M bromide or thiocyanate are added.

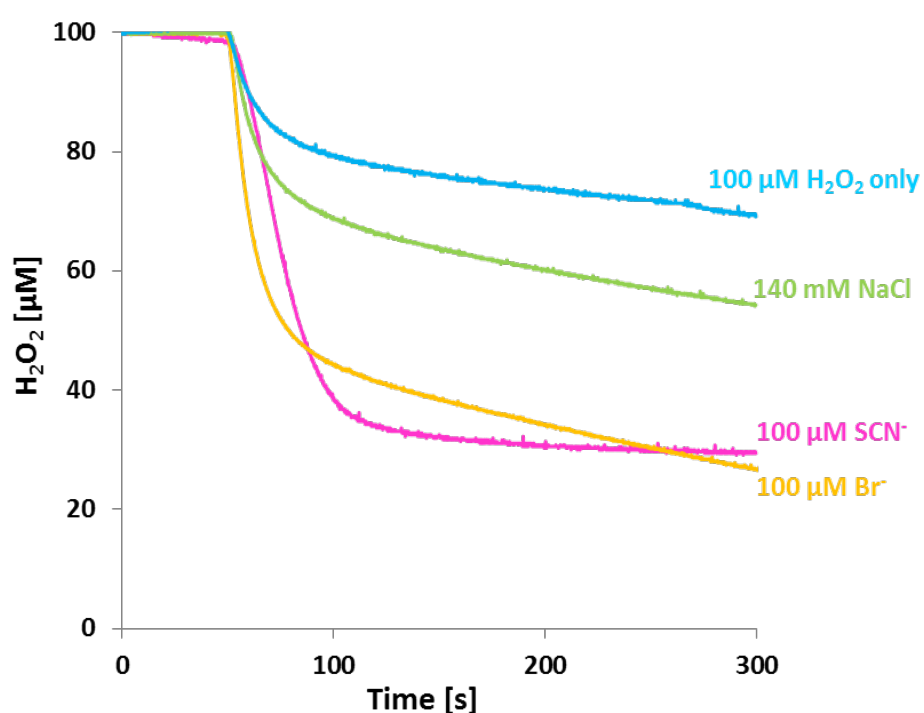


Figure 4-25: Hydrogen peroxide consumed by Con 4N in the presence of chloride, thiocyanate and bromide: Measurements were performed with 500 nM enzyme in 100 mM phosphate buffer, pH 7.

Furthermore, the peroxidase substrates urate, tyrosine and ascorbate were also tested in the same system.

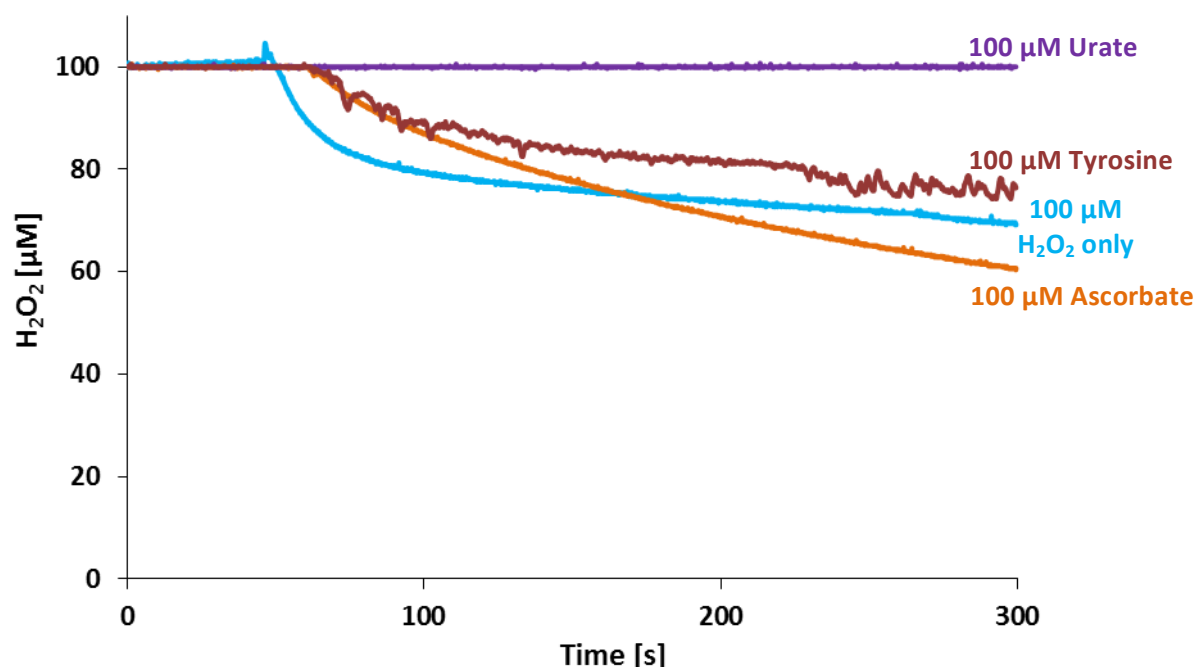


Figure 4-26: Hydrogen peroxide consumed by Con 4N in the presence of urate, tyrosine and ascorbate:
Measurements were performed with 500 nM enzyme in 100 mM phosphate buffer, pH 7.

Urate inhibited the turnover of hydrogen peroxide by Con 4N completely. The substrates tyrosine and ascorbate reduced the consumption of hydrogen peroxide and the initial velocity of all these reactions is slower than with hydrogen peroxide only (Figure 4-26).

This is the first time that the turnover of these substrates was shown directly. Once more it was confirmed that bromide is the physiological substrate for hsPxd1. But the results also demonstrate that thiocyanate is a good electron donor and that these two substrates would compete for hsPxd01 in a physiological setting.

Next the concentration dependency was recorded for these two substrates. Figure 4-27 shows the curves for varying bromide concentrations from 20 – 200 μM. It is obvious that the reaction is concentration dependent. From the initial slopes of these curves the kinetic constants were again determined.

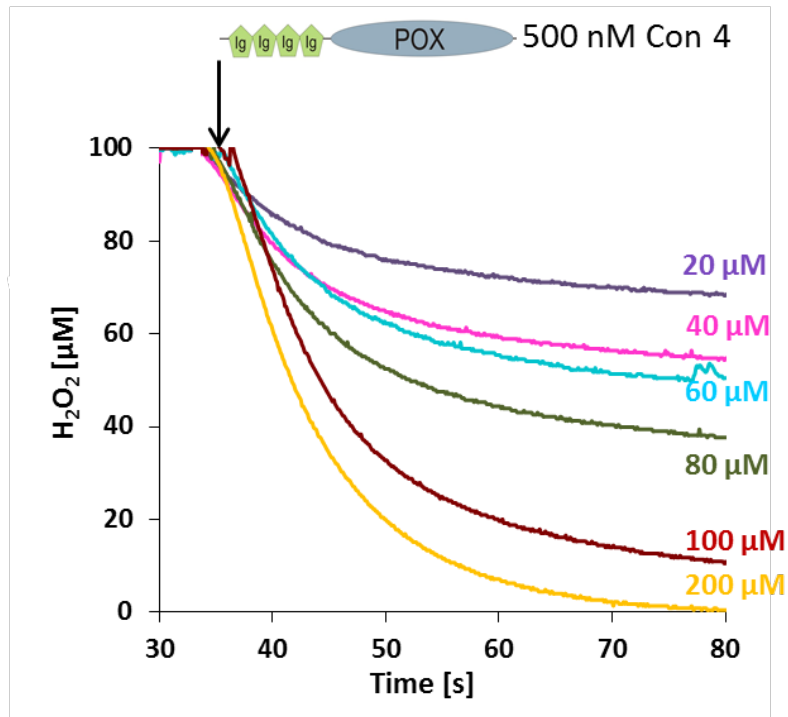


Figure 4-27: Varying Br^- concentrations showed that the reaction is concentration-dependent:
initial 100 μM H_2O_2 in 100 mM phosphate buffer pH 7, 500 nM enzyme was added to start the reaction

The **Hanes plot** is another method for linearization of the Michaelis-Menten equation (see also Lineweaver-Burk plot in chapter 4.3.6.2). Figure 4-28 represents the basic principle of this method.

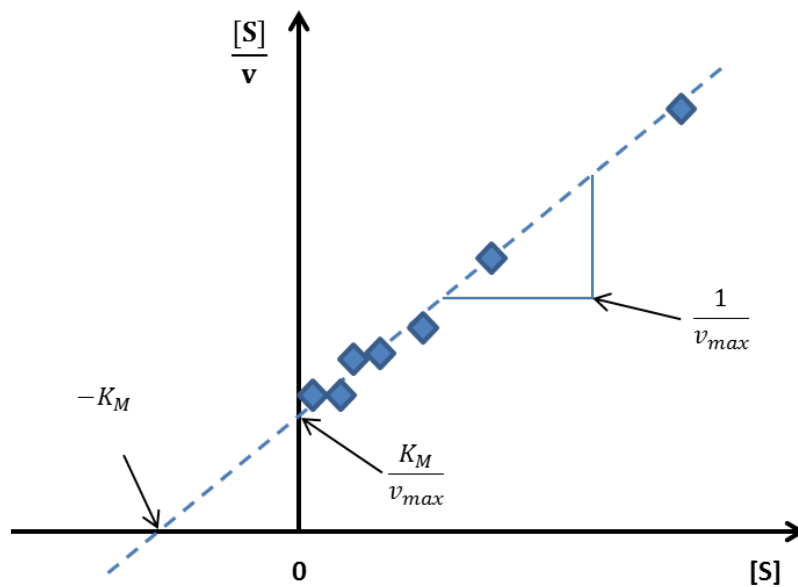


Figure 4-28: Scheme of a Hanes plot

By plotting the determined values and adding a trendline the linear equation was determined – as seen in Figure 4-29 – to be:

$$y = 0.0919 * x + 8.0385$$

From the scheme depicted in Figure 4-28 the following parameters were determined:

If $x = 0 \rightarrow y = \frac{K_M}{v_{max}}$

And $y = 0 \rightarrow x = -K_M$

Table 4-8 summarizes the obtained values from the plot in Figure 4-29 which are higher than those obtained from the indirect activity assay:

Table 4-8: Summary of the kinetic constants

v_{max}	652.83 $\mu\text{M} \cdot \text{min}^{-1}$
K_M	87.46 μM
k_{cat}	1305.66 min^{-1}
k_{cat} / K_M	$1.49 \cdot 10^7 \text{ M}^{-1} \cdot \text{min}^{-1}$

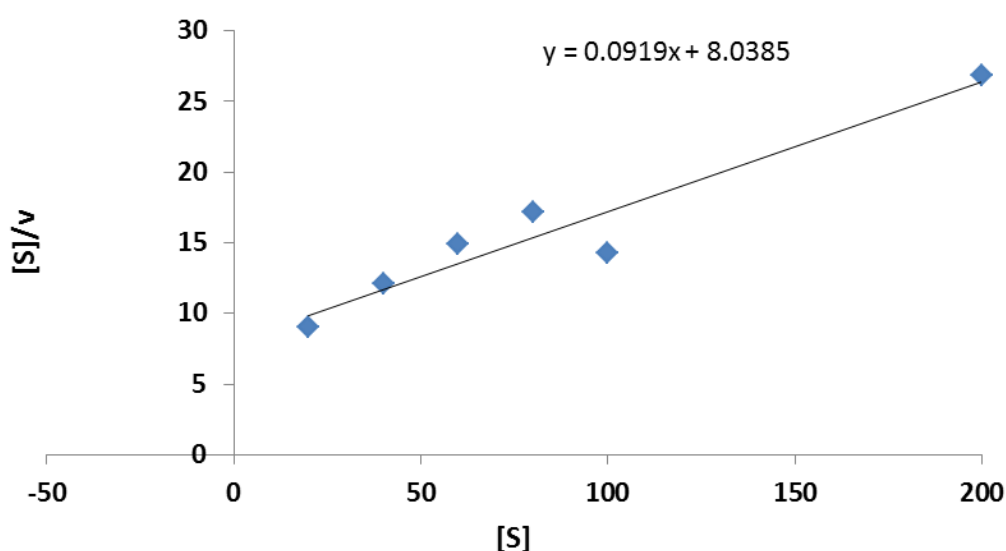


Figure 4-29: Hanes plot for the determination of kinetic constants: the system contained 500 nM Con 4N, 100 mM phosphate buffer, pH 7, 100 μM H_2O_2 and varying concentrations of Br^- (20-200 μM).

The same procedure was performed for thiocyanate: Concentrations from 20 to 200 μM were added. Figure 4-30 shows the resulting curves. The maximum seems to be reached at a concentration of 60 μM from there on it seems as if the reaction gets inhibited by thiocyanate. This resulted in a Hanes plot with a negative intercept (see Figure 4-31) which can't be solved. The reason is not clear at the moment but this at least explains why it was not possible to obtain reasonable results for the TNB assay as well (Chapter 4.3.6.2).

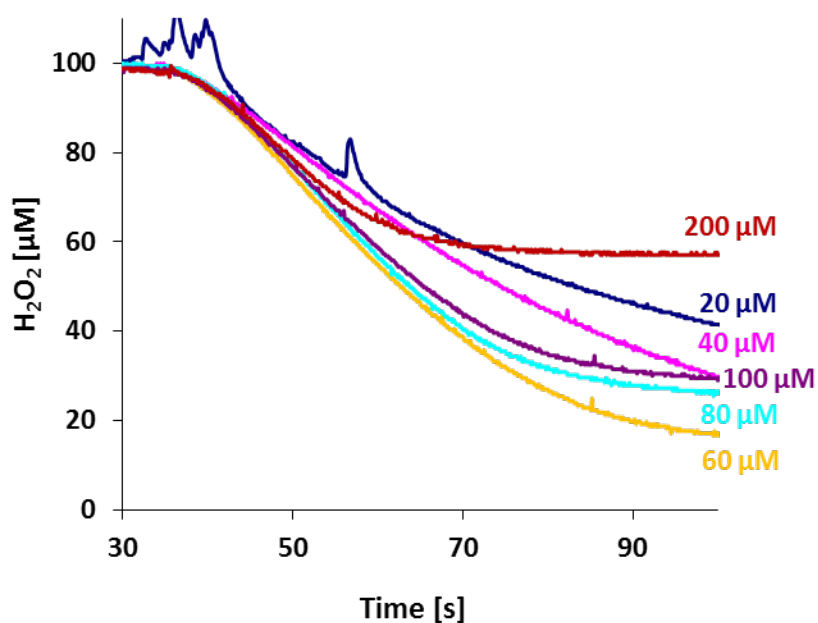


Figure 4-30: Varying concentrations of thiocyanate showed an inhibition of the reaction at higher concentrations: initial 100 μM H_2O_2 in 100 mM phosphate buffer pH 7, 500 nM enzyme was added to start the reaction.

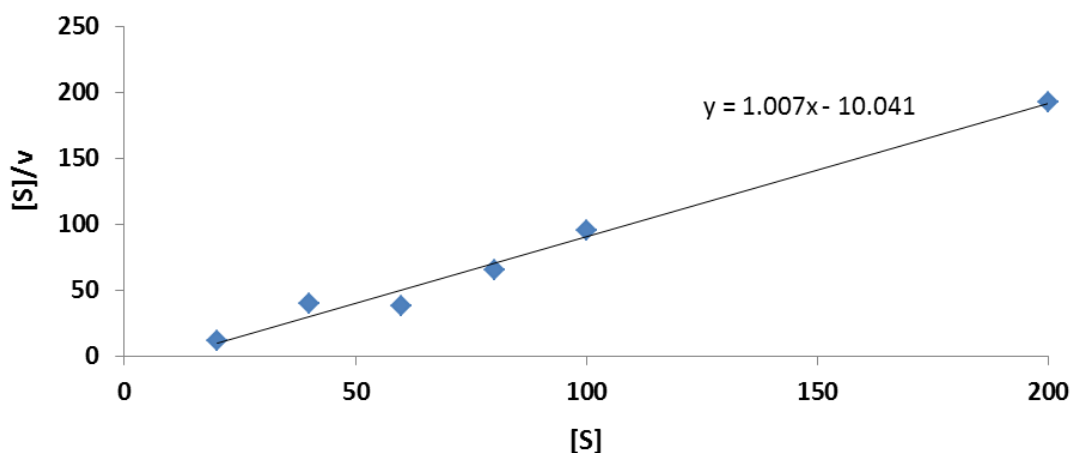


Figure 4-31: Hanes Plot for varying thiocyanate concentrations: determination of kinetic constants was not possible. The system contained 500 nM Con 4N, 100 μM H_2O_2 and varying concentrations of SCN^- (20-200 μM) in 100 mM phosphate buffer, pH 7.

4.3.8 Pre-steady state kinetics

4.3.8.1 *Compound I formation*

When the ferric enzyme reacts with hydrogen peroxide, Compound I is formed which then oxidizes halogens to their respective hypohalous acids.

The formation of Compound I and the subsequent formation of Compound II is shown in Figure 4-32. The drop of the absorbance maximum at 412 nm over time depicts Compound I formation and is completed within the first 200-300 ms. The subsequent shift of the Compound I spectrum resulting in the appearance of a shoulder at 433 nm could either be the decay of Compound I or formation of Compound II. As shown below Con 4N forms Compound I and Compound II in a concentration dependent manner.

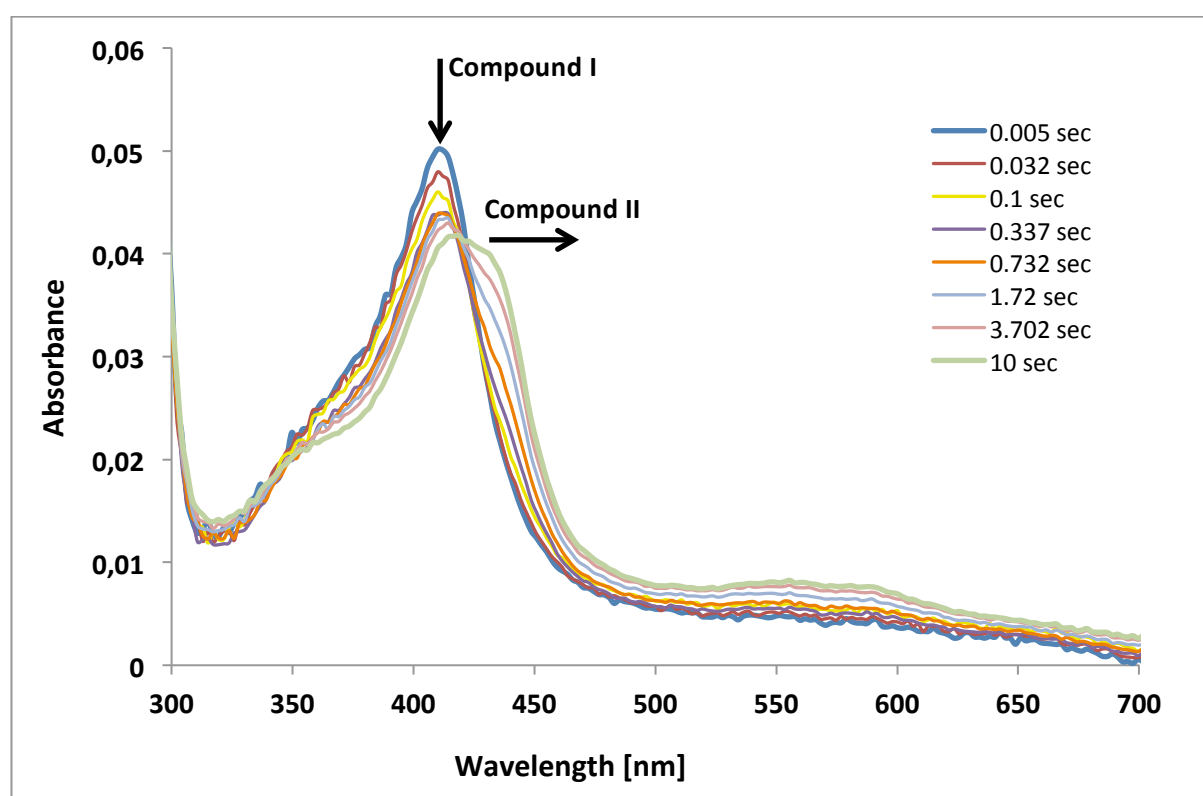


Figure 4-32: Spectra of the reaction of Con 4N with hydrogen peroxide: the blue line represents the start of the reaction. Over time the absorbance decreased and a shift (green line) is visible at the end of the reaction after 10 seconds. The concentrations in the measuring cell were: 500 nM Con 4N and varying H_2O_2 concentrations from 2.5 μM to 10 μM in 100 mM phosphate buffer, pH 7.4.

The reaction further was followed at a single wavelength (412 nm) as depicted in Figure 4-33. The resulting single trace curve was used to determine the k_{obs} values at varying hydrogen peroxide concentrations. Therefore the initial fast reaction was fitted double exponentially using the *Pro Data Viewer* programme.

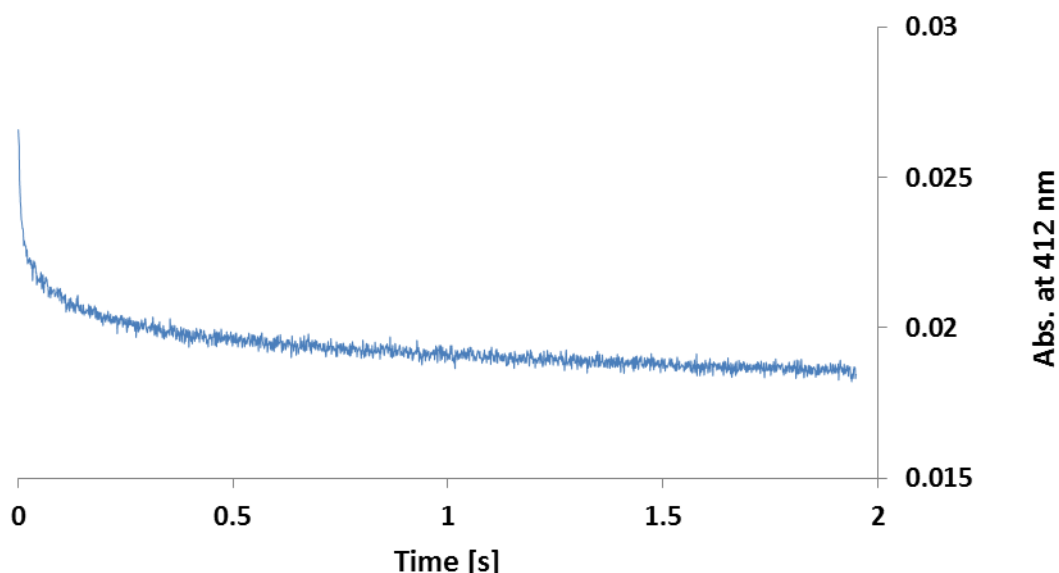


Figure 4-33: Time trace of the reaction of Con 4N with hydrogen peroxide: the initial fast reaction represents the formation of Compound I of Con 4N which is then followed by the much slower formation of Compound II.

The resulting k_{obs} values of the first part of the double exponential fit were plotted against the various hydrogen peroxide concentrations. Again a trendline was placed through those points and the resulting linear equation was used to determine the k_{app} value from the slope representing the rate constant of Compound I formation.

As seen in Figure 4-34 the determined linear equation was:

$$y = 16.165 * x + 18.195$$

Whereby the slope equals the apparent bimolecular rate constant k_{app} which resulted to be $16.165 \mu\text{M}^{-1}\text{s}^{-1}$ or:

$$k_{app} = 1.6 \cdot 10^7 \text{ M}^{-1}\text{s}^{-1}$$

It is clearly visible that the reaction is concentration dependent.

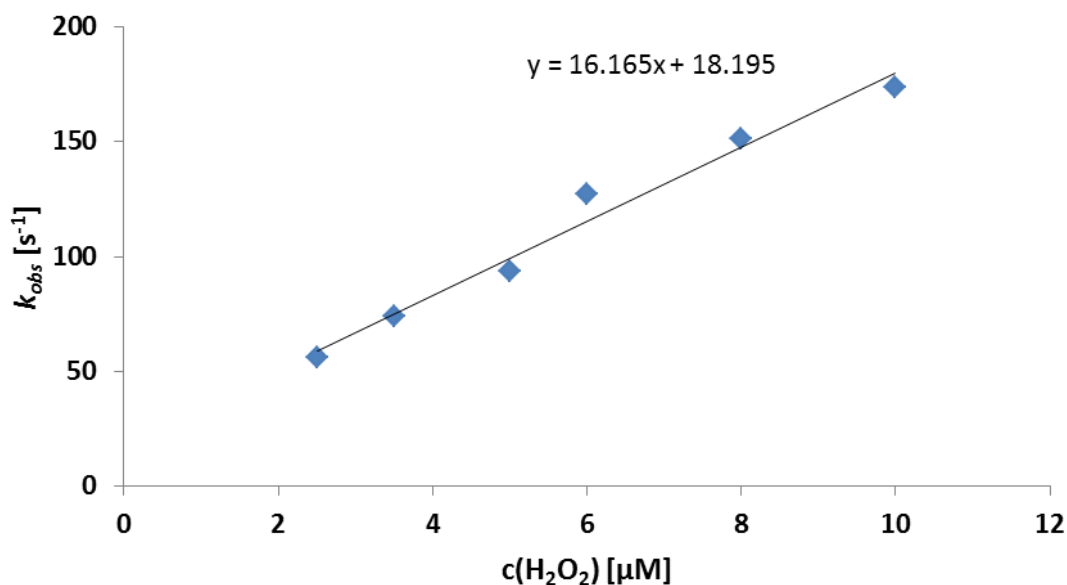


Figure 4-34: Determination of the k_{app} value of Compound I formation by plotting the k_{obs} value vs. the distinct hydrogen peroxide concentrations at pH 7.4 (enzyme diluted in 100 mM phosphate buffer).

Also the second part of the double exponential fit of the curves as shown in Figure 4-33 were used to determine k_{obs} values which were plotted against the distinct H_2O_2 concentrations. Again the slope of the resulting trendline revealed a k_{app} value of the second part of the reaction which was determined to be:

$$k_{app} = 6.1 \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$$

and represents Compound II formation. As this part was also concentration dependent it was concluded that Compound II was formed from Compound I. This is interesting because the Compound I of LPO – the enzyme to which peroxidasin is compared often based on the high amino acid sequence homology – doesn't form Compound II but decays by forming Compound I*, which is an intermediate where Compound I oxidizes one of the amino acid residues close to the active site resulting in a protein radical. Peroxidasin, or at least Con 4N, seems to behave more like MPO than LPO in this case, as MPO is able to use H_2O_2 as one-electron donor of Compound I, thereby producing Compound II (see reaction scheme in Figure 1-7, Reaction 3b).

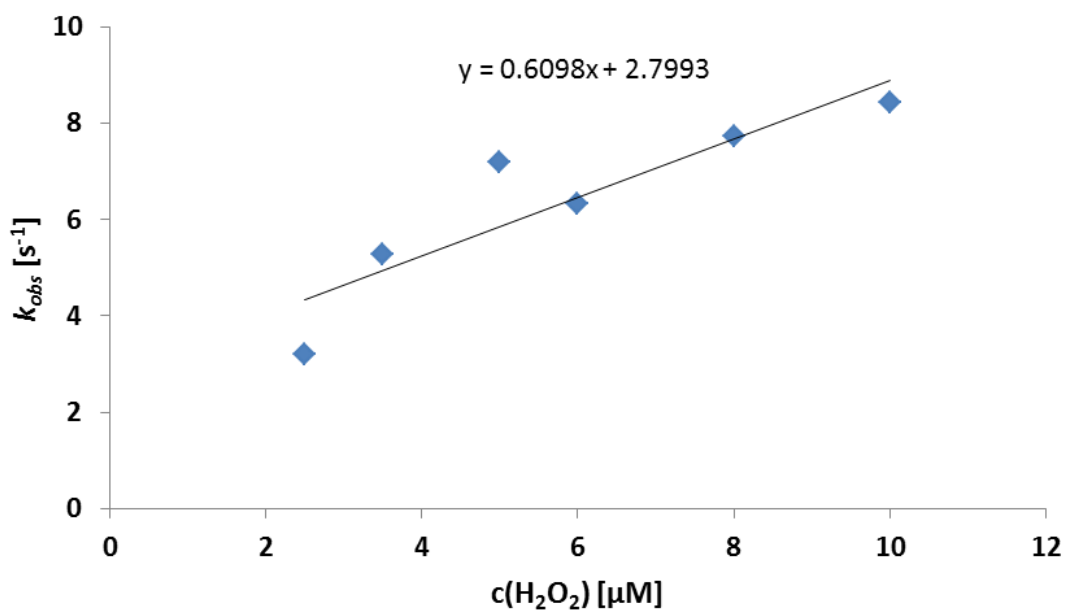


Figure 4-35: Determination of the k_{app} value of the reaction between Compound I and H_2O_2 by plotting the k_{obs} values vs. the distinct hydrogen peroxide concentrations in 100 mM phosphate buffer pH 7.4.

4.3.8.2 Cyanide binding:

Cyanide binds tightly to the Fe(III) of the heme in the active site of peroxidases. This results in the conversion of the high-spin to a low-spin state, which further alters the absorption maximum of the Soret peak. This process is an indication for the accessibility to the active site.

Figure 4-36 shows the UV-vis spectra of 1 μM Con 4N treated with 1 mM NaCN. The heme peak was red-shift from 412 nm to 431 nm in the course of 10 seconds and the absorbance was also decreasing.

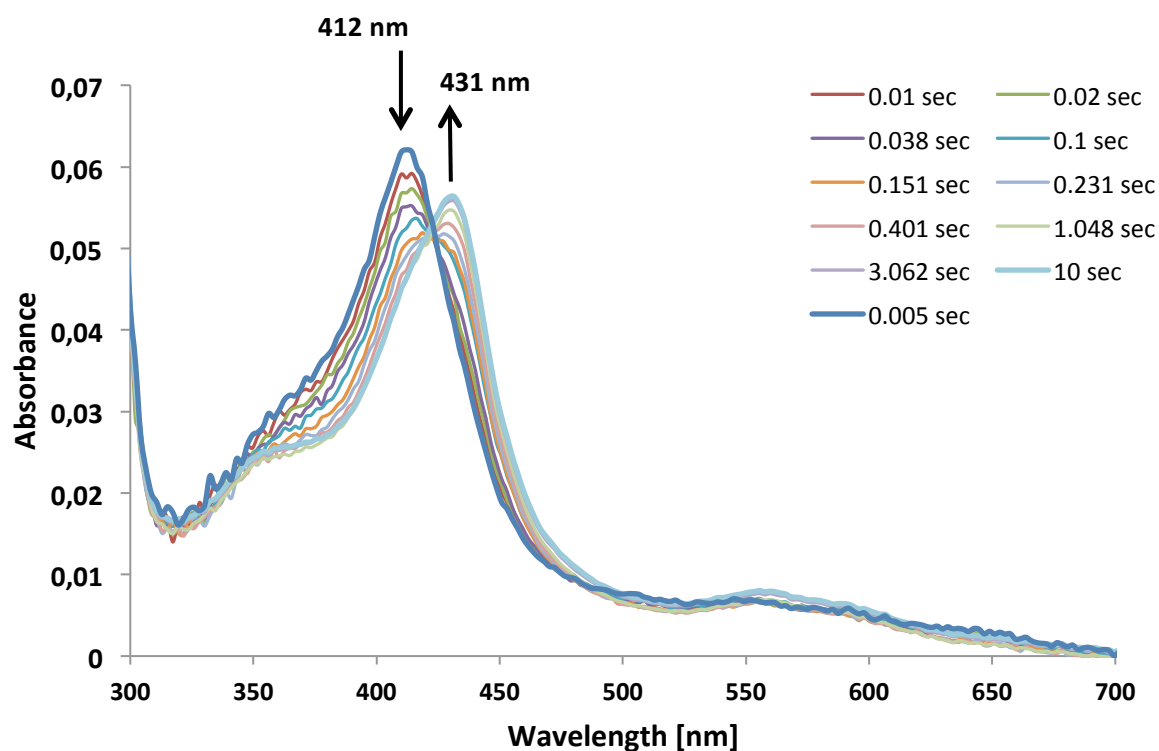


Figure 4-36: UV-Vis spectra of Con 4N reacting with cyanide resulting in a shift of the Soret peak over time

Further, the time trace at 434 nm is exemplarily shown in Figure 4-37. This trace was again fitted double exponentially. The k_{obs} values determined from the first steep part of the curves were plotted against cyanide concentrations, as shown in Figure 4-38. From the resulting linear equation k_{on} and k_{off} could be determined. The slope – in this case 0.5362 – equals k_{on} and the intercept conforms to k_{off} . These two values were used to determine the dissociation constant K_D which gives information about the affinity of cyanide to the heme:

$$K_D = \frac{k_{off}}{k_{on}}$$

The values are summarized in Table 4-9.

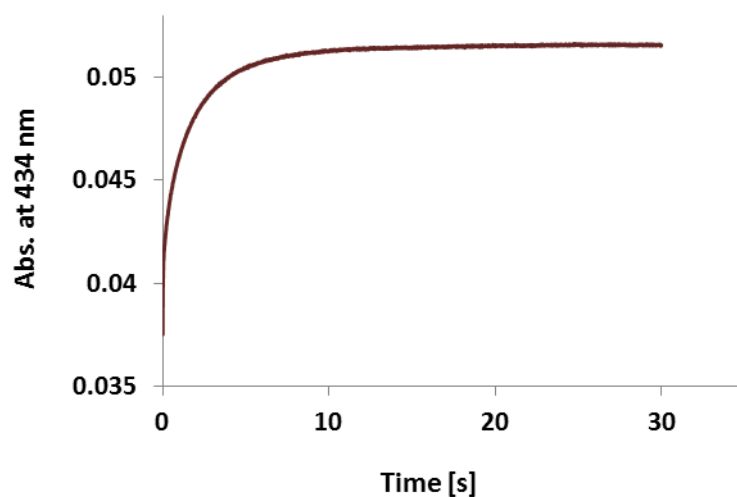


Figure 4-37: Time trace of cyanide binding to Con 4N

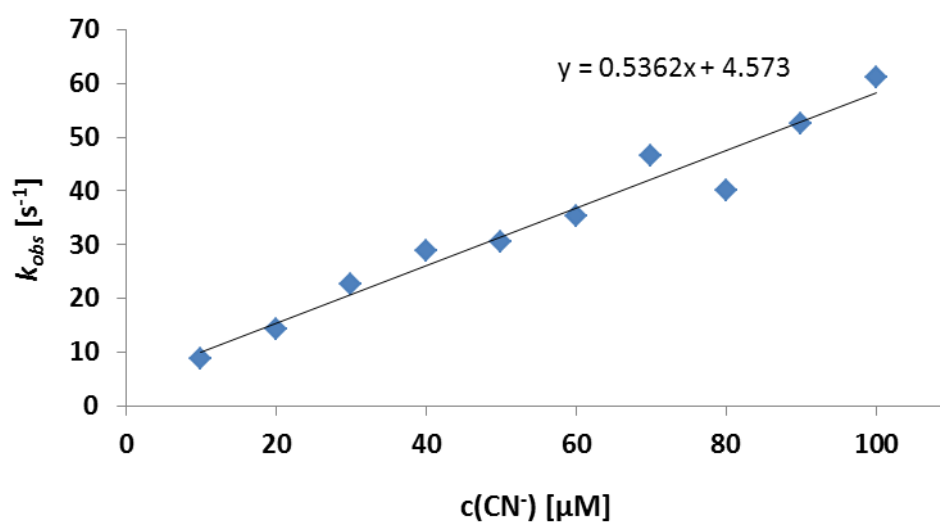


Figure 4-38: k_{obs} values of different cyanide concentrations to determine k_{on} , k_{off} and K_D

Table 4-9: Summary of values determined for cyanide binding

k_{on}	0.54 $\mu\text{M}^{-1}\text{s}^{-1}$
k_{off}	4.58 s^{-1}
K_D	8.5 μM

A K_D value of 8.5 μM is also a typical value for peroxidases. The k_{on} value reflects a good accessibility of the active site.

4.4 Structural investigations

4.4.1 X-ray crystallography

As described above, many different screens and conditions were tested to produce crystals. This would resolve the structure of the active site and reveal the organization of the different domains. So far, all but one of resulting crystals turned out to be salt crystals. The protein one didn't scatter enough to give results. There are different methods which still could be tested to promote crystallization like *in situ* proteolysis, deglycosylation and others but these are time consuming and couldn't be performed in the course of this thesis.

4.4.2 Preliminary SAXS data results

Because of the poor crystallisation results, SAXS measurements of Con 5N and Con 4N were performed meanwhile. These resulting data should give an insight in the arrangement of the peroxidase domain and the four Ig domains relative to each other. One has to be aware that these results are not as precise due to lower resolution as crystallography data but they can provide an approach to the overall organization of a protein.

The complex data set from SAXS measurements were kindly processed and evaluated by Dr. Rupert Tscheließnig. Therefore the so-called SAXS profile had to be transformed into an overall shape. In this shape the model of the peroxidase domain of hsPxd01 based on the X-ray structure of goat LPO as depicted in blue (Figure 4-39) was incorporated. The red part represents a part of the POX domain named *propeptide homologue* as its sequence is similar to the propeptide of human peroxidases but it is not cleaved off. This part could not be modelled as it's not present in mature LPO. The propeptide of LPO is proteolytically removed during maturation. The propeptide homologue seems to protrude from the peroxidase model and is therefore believed to be a region of high flexibility.

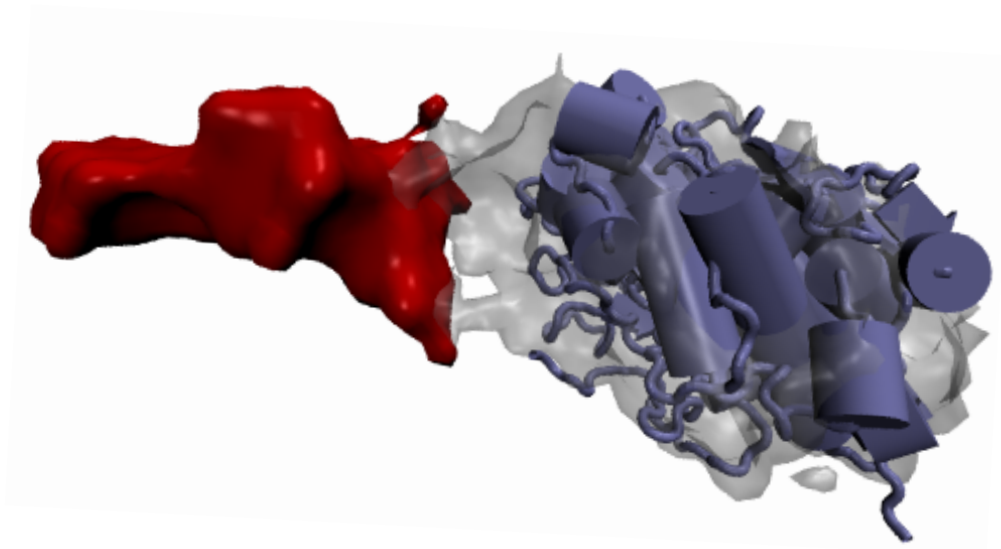


Figure 4-39: SAXS result for the peroxidase domain (Con 5N) modelled on gLPO.

The second structure below (Figure 4-40) resulted from the SAXS data of Con 4N.

The overall shape of the Con 5N model was used as a template. As can be seen in Figure 4-40 the Ig₃ and Ig₄ domains seem to be associated with the peroxidase domain. The whole Ig region appears to be a region of high flexibility.

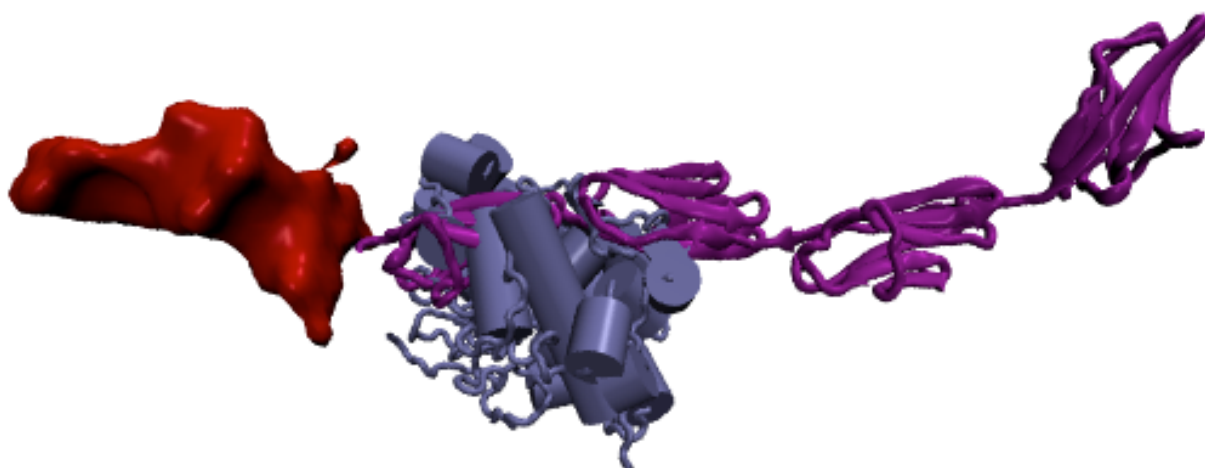


Figure 4-40: SAXS result for the peroxidase domain and the four Ig domains (Con 4N)

5 Conclusions

Although human peroxidasin 1 is thought to play an important role in human physiology, comparatively little is known about this multidomain heme enzyme. One reason therefore is, that this 504 kDa large homotrimer could not be expressed recombinantly in sufficient amounts so far. Hence, considerations about working with shorter variants of this protein were made, which further provided the basis for this thesis.

Consequently, different variants of human peroxidasin 1 were designed and cloned by omitting distinct domains of the full length enzyme. Of course, the catalytically active peroxidase domain was part of each and every one of them. Expression in HEK293-6E cells and subsequent purification was performed to further conduct biochemical characterisation studies as well as structural investigations.

Therefore, successful cloning of ten different truncated variants of human peroxidasin 1 was performed using new cloning strategies. All of them were tested in the course of small scale experiments and by now, six of them were also expressed in large scale. Expression in HEK 293-6E cells was proven to be a reliable and comparatively fast way to produce the constructs in mammalian cells.

The purification of the harvested cell supernatant could be significantly improved for one construct (Con 4N). An additional incubation step showed to enhance covalent link formation which further leads to a more active enzyme. This assumption could be reproduced and tests performed with this new, enhanced variant showed higher activities than peroxidasin showed ever before, disproving the assumption that the active site might not be accessible. Purification numbers and yield varied for the different constructs. The improved Con 4N variant could be gained in large amounts of about 19 mg/L cell supernatant, although the purity number of averagely 0.4 was not as good as for Con 5N (0.6).

Furthermore, SDS-PAGE and Western blotting of constructs could reveal the exact sites responsible for trimerization of human peroxidasin 1. It could be clearly shown that the linker region connecting peroxidase domain and VWC domain – consisting of cysteine residues and an amphipathic helical sequence – is responsible for oligomerization and that the VWC domain is not involved at all.

Widespread biophysical and biochemical characterization tools were used to systematically characterize the truncated variants. Due to amino acid sequence homologies it was concluded that peroxidasin should behave similar to the related chordata peroxidase LPO to which it therefore was compared often. As collateral performed expression and purification were improved just after a few months the methods were performed with the “poorer”, as well as with the improved samples of Con 4N. Some experiments were carried out with different constructs to compare the outcomes but the majority is based on the improved Con 4N. The similarity of Con 4N to LPO could for example be shown by circular dichroism measurements of the heme environment or UV-Vis spectra revealing a Soret maximum of 412 nm and typical bands in the visible region. Furthermore the activity measurements performed revealed LPO like behaviour.

It could be confirmed that bromide is the physiological substrate of hsPxd01 Con 4N by showing – for the first time directly – the turnover of this substrate and subsequent determination of the K_M value as well as the k_{cat} and v_{max} value. Moreover it was shown that thiocyanate seems to be a very good substrate too, which means that both bromide and thiocyanate will be turned over in a physiological setting competing for the enzyme. In contrast to many other peroxidases, classical peroxidase substrates as ABTS and guaiacol do not seem to be converted as good by peroxidasin resulting in very low specific activity of < 1 U/mg. Halogenation activity in contrast lead to far higher values for both, bromide and thiocyanate.

Pre-steady state kinetics using stopped-flow spectroscopy was performed to investigate the redox intermediates of Con4N when reacting with hydrogen peroxide. By determination of the rate constant k_{app} it could be confirmed that peroxidasin Con 4N behaves similar to other peroxidases. Interestingly hsPxd01 Con 4N Compound I reacts with hydrogen peroxide to form Compound II (ferric protein $\rightarrow$ Compound I $\rightarrow$ Compound II). LPO Compound I cannot be reduced with hydrogen peroxide to Compound II. So this result is a striking difference to LPO and the behaviour is MPO like in this respect.

Further, binding of cyanide could be used to determine a dissociation constant K_D of 8.5 μ M, which once more indicates the accessibility of the active site and showed expected peroxidase behaviour.

Moreover, structural investigations were performed. Clear insights into the 3D structure of various constructs would be given by crystallography but as experiments so far did not produce suitable crystals more work on this would be needed. Preliminary SAXS data was generated revealing an overall structure of hsPxd Con5N and Con4N, which suggested regions of high flexibility.

6 References

- Abreu-Velez, A. M., and M. S. Howard, 2012, Collagen IV in Normal Skin and in Pathological Processes: *N Am J Med Sci*, v. 4, p. 1-8.
- Baldi, L., N. Muller, S. Picasso, R. Jacquet, P. Girard, H. P. Thanh, E. Derow, and F. M. Wurm, 2005, Transient gene expression in suspension HEK-293 cells: application to large-scale protein production: *Biotechnol Prog*, v. 21, p. 148-53.
- Bhave, G., C. F. Cummings, R. M. Vanacore, C. Kumagai-Cresse, I. A. Ero-Tolliver, M. Rafi, J. S. Kang, V. Pedchenko, L. I. Fessler, J. H. Fessler, and B. G. Hudson, 2012, Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis: *Nat Chem Biol*, v. 8, p. 784-90.
- Birnboim, H. C., and J. Doly, 1979, A rapid alkaline extraction procedure for screening recombinant plasmid DNA: *Nucleic Acids Res*, v. 7, p. 1513-23.
- Boldon, L., F. Laliberte, and L. Liu, 2015, Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application: *Nano Rev*, v. 6, p. 25661.
- Cheng, G., J. C. Salerno, Z. Cao, P. J. Pagano, and J. D. Lambeth, 2008, Identification and characterization of VPO1, a new animal heme-containing peroxidase: *Free Radic Biol Med*, v. 45, p. 1682-94.
- Furtmüller, P. G., M. Zederbauer, W. Jantschko, J. Helm, M. Bogner, C. Jakopitsch, and C. Obinger, 2006, Active site structure and catalytic mechanisms of human peroxidases: *Arch Biochem Biophys*, v. 445, p. 199-213.
- Gorrec, F., 2009, The MORPHEUS protein crystallization screen: *J Appl Crystallogr*, v. 42, p. 1035-1042.
- Hellmark, T., and M. Segelmark, 2014, Diagnosis and classification of Goodpasture's disease (anti-GBM): *J Autoimmun*, v. 48-49, p. 108-12.
- Horikoshi, N., J. Cong, N. Kley, and T. Shenk, 1999, Isolation of differentially expressed cDNAs from p53-dependent apoptotic cells: activation of the human homologue of the *Drosophila* peroxidasin gene: *Biochem Biophys Res Commun*, v. 261, p. 864-9.
- Jantschko, W., P. G. Furtmüller, M. Zederbauer, C. Jakopitsch, and C. Obinger, 2004, Kinetics of oxygen binding to ferrous myeloperoxidase: *Arch Biochem Biophys*, v. 426, p. 91-7.
- Longo, P. A., J. M. Kavran, M. S. Kim, and D. J. Leahy, 2013, Transient mammalian cell transfection with polyethylenimine (PEI): *Methods Enzymol*, v. 529, p. 227-40.

- Lázár, E., Z. Péterfi, G. Sirokmány, H. A. Kovács, E. Klement, K. F. Medzihradsky, and M. Geiszt, 2015, Structure-function analysis of peroxidasin provides insight into the mechanism of collagen IV crosslinking: *Free Radic Biol Med*, v. 83, p. 273-82.
- McCall, A. S., C. F. Cummings, G. Bhavé, R. Vanacore, A. Page-McCaw, and B. G. Hudson, 2014, Bromine is an essential trace element for assembly of collagen IV scaffolds in tissue development and architecture: *Cell*, v. 157, p. 1380-92.
- Nelson, R. E., L. I. Fessler, Y. Takagi, B. Blumberg, D. R. Keene, P. F. Olson, C. G. Parker, and J. H. Fessler, 1994, Peroxidasin: a novel enzyme-matrix protein of *Drosophila* development: *EMBO J*, v. 13, p. 3438-47.
- Péterfi, Z., A. Donkó, A. Orient, A. Sum, A. Prókai, B. Molnár, Z. Veréb, E. Rajnavölgyi, K. J. Kovács, V. Müller, A. J. Szabó, and M. Geiszt, 2009, Peroxidasin is secreted and incorporated into the extracellular matrix of myofibroblasts and fibrotic kidney: *Am J Pathol*, v. 175, p. 725-35.
- Péterfi, Z., Z. E. Tóth, H. A. Kovács, E. Lázár, A. Sum, A. Donkó, G. Sirokmány, A. M. Shah, and M. Geiszt, 2014, Peroxidasin-like protein: a novel peroxidase homologue in the human heart: *Cardiovasc Res*, v. 101, p. 393-9.
- Soudi, M., M. Paumann-Page, C. Delporte, K. F. Pirker, M. Bellei, E. Edenhofer, G. Stadlmayr, G. Battistuzzi, K. Z. Boudjeltia, P. G. Furtmüller, P. Van Antwerpen, and C. Obinger, 2015, Multidomain human peroxidasin 1 is a highly glycosylated and stable homotrimeric high spin ferric peroxidase: *J Biol Chem*, v. 290, p. 10876-90.
- Soudi, M., M. Zamocky, C. Jakopitsch, P. G. Furtmüller, and C. Obinger, 2012, Molecular evolution, structure, and function of peroxidasins: *Chem Biodivers*, v. 9, p. 1776-93.
- Thomas, P., and T. G. Smart, 2005, HEK293 cell line: a vehicle for the expression of recombinant proteins: *J Pharmacol Toxicol Methods*, v. 51, p. 187-200.
- Wlodawer, A., W. Minor, Z. Dauter, and M. Jaskolski, 2008, Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures: *FEBS J*, v. 275, p. 1-21.
- Wurm, F. M., 2004, Production of recombinant protein therapeutics in cultivated mammalian cells: *Nat Biotechnol*, v. 22, p. 1393-8.
- Zamocky, M., C. Jakopitsch, P. G. Furtmüller, C. Dunand, and C. Obinger, 2008, The peroxidase-cyclooxygenase superfamily: Reconstructed evolution of critical enzymes of the innate immune system: *Proteins*, v. 72, p. 589-605.
- Zederbauer, M., P. G. Furtmüller, S. Brogioni, C. Jakopitsch, G. Smulevich, and C. Obinger, 2007, Heme to protein linkages in mammalian peroxidases: impact on spectroscopic, redox and catalytic properties: *Nat Prod Rep*, v. 24, p. 571-84.

Zámocký, M., and C. Obinger, 2010, Molecular Phylogeny of Heme Peroxidases. Biocatalysis Based on Heme Peroxidases: Peroxidases as Potential Industrial Biocatalysts, 7-35

Zámocký, M., S. Hofbauer, I. Schaffner, B. Gasselhuber, A. Nicolussi, M. Soudi, K. F. Pirker, P. G. Furtmüller, and C. Obinger, 2015, Independent evolution of four heme peroxidase superfamilies: Arch Biochem Biophys, v. 574, p. 108-19.