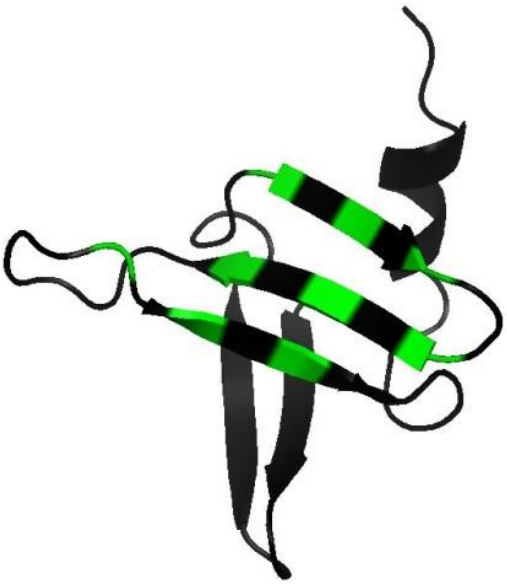
UNIVERSITY OF NATURAL RESOURCES AND LIFE SCIENCES, VIENNA
DEPARTMENT OF CHEMISTRY
DIVISION OF BIOCHEMISTRY
PROTEIN BIOCHEMISTRY GROUP

ENGINEERING MINI-BINDERS FOR MOLECULAR RECOGNITION OF HUMAN PEROXIDASIN



MASTER THESIS

SUBMITTED BY
NIKOLAUS F. KIENZL

VIENNA, 2018

SUPERVISOR:
CHRISTIAN OBINGER, UNIV.PROF. MAG. DR.RER.NAT.
CO-SUPERVISOR:
MICHAEL TRAXLMAYR, DIPL.-ING. DR.

University of Natural Resources
and Life Sciences, Vienna

Φύσις κρύπτεσθαι φιλεῖ.

Die Natur liebt es, sich zu verbergen.

Heraklit

Acknowledgements

First, I want to say thank you for the great chance working in your research group, being a member of the team of Bio B. You always provided me with your advice and expertise. It was a great experience sharing my time in- and outside the lab with you – always in a very pleasant and positive atmosphere.

Special thanks to Christian Obinger who gave me the opportunity to perform my master thesis in his team. Thank you to my excellent supervisor Michael Traxlmayr. With patience and a lot of expertise, he supported me during my master thesis. Thank you.

Furthermore, I want to mention and to say thank you to my closest friends Oliver Törzsök, Philipp Schmuck, Clemens Petrasch, Christian Hrdina and Jakob Hellinger. They supported me with critical input and questions - from a scientific and a non-scientific point of view.

Last but not least many thanks to my family - especially my mother Eveline Kienzl, my brother Philipp Kienzl and my grandfather, Robert Spielvogl.

Abstract

The aim of this thesis was to engineer mini-binders for the molecular recognition of human peroxidasin 1, which plays a critical role in the consolidation of the basement membrane, a specialized form of the extracellular matrix. Peroxidasin, a homotrimeric, highly glycosylated iron protein with a covalently bound ferric high spin heme, uses bromide as a cofactor for the formation of sulfilimine (S=N) cross-links between two adjacent collagen IV protomers. Until now, no crystal structure could be obtained, partly due to the low expression rate of peroxidasin. Further intricacies are the enormous size of human peroxidasin 1, glycosylations as well as flexible domains. However, adding a stable interaction partner during the crystallization process might facilitate crystallization of this large multidomain protein. Moreover, having access to such a peroxidasin-binding molecule would also enable purification of endogenous peroxidasin by affinity chromatography. Therefore, a reduced-charge variant of the small DNA-binding protein Sso7d from the hyperthermophilic archaeon *Sulfolobus solfataricus* was engineered for binding to human peroxidasin 1 and its truncated versions. Two yeast display libraries were used for the selection process. Finally, we obtained 11 binders belonging to four families recognizing at least two different epitopes on peroxidasin with binding affinities down to the single-digit nanomolar range. These binders may facilitate the purification process of endogenous peroxidasin and could further function as crystallization chaperones in order to elucidate the structure of peroxidasin. Future applications of those peroxidasin binders could include the use as a detection reagent in ELISA assays, Western blots, flow cytometry or microscopy.

Zusammenfassung

Das Ziel dieser Arbeit war die Generierung kleiner Bindermoleküle für die molekulare Erkennung von humanem Peroxidasin, das eine essentielle Rolle in der Festigung der Basalmembran, einer spezialisierten Form der extrazellulären Matrix, spielt. Peroxidasin, ein homotrimeres, stark glykosyliertes Eisenprotein mit kovalent gebundenem high-Spin Häm, verwendet Bromid als Kofaktor für die Bildung der Sulfiliminquerverbindung (S=N) zweier benachbarter Kollagen IV Protomere. Da die Expression und Aufreinigung von Peroxidasin zu geringe Mengen an Protein liefert, konnte noch keine Kristallstruktur aufgeklärt werden. Weitere Schwierigkeiten stellen die Größe, die Glykosylierungen sowie die flexiblen Domänen von humanem Peroxidasin 1 dar. Das Hinzufügen eines stabilen Interaktionspartners während des Kristallisationsprozesses könnte diesen für das große Multidomänenprotein vereinfachen. Des Weiteren würde ein derartiges Bindermolekül die Aufreinigung endogenen Peroxidasins mittels Affinitätschromatographie ermöglichen. Daher wurden auf Basis einer ladungsreduzierten Variante des kleinen DNA-bindenden Proteins Sso7d, das aus dem hyperthermophilen Archaeon *Sulfolobus solfataricus* stammt, hoch-affine Binder gegen Peroxidasin und seine verkürzten Varianten konstruiert. Zwei Hefe-Display-Bibliotheken wurden für den Selektionsprozess verwendet. Schließlich erhielten wir elf Binder (in vier Familien zuordenbar), die an mindestens zwei überlappende Epitope auf Peroxidasin mit Affinitäten bis in den einstelligen Nanomolarbereich binden. Diese Binder könnten den Aufreinigungsprozess von endogenem Peroxidasin ermöglichen und könnten weiters als Kristallisationschaperone fungieren, um die Struktur des Peroxidasins aufzuklären. Der Einsatz als Detektionsreagenz in ELISA assays, Western blots, Durchflusszytometriemessungen oder in der Mikroskopie, könnte zusätzliche zukünftige Anwendungen der Peroxidasinbinder darstellen.

1. Introduction

1.1. Heme Peroxidase Families

Heme-containing peroxidases like myeloperoxidase (MPO), eosinophile peroxidase (EPO), lactoperoxidase (LPO) and thyroid peroxidase (TPO) are ubiquitous oxidoreductases able to cleave peroxidic bonds heterolytically. They use heme *b* or posttranslationally modified heme as redox cofactor. The reaction mechanism common to all peroxidases is the reduction of hydrogen peroxide to water in either a one- or two-electron oxidation. Hydrogen peroxide turnover can lead to different reactions:

(1) $H_2O_2 + 2AH_2 \rightarrow 2H_2O + 2{\cdot}AH$
(2) $H_2O_2 + X^- + H^+ \rightarrow H_2O + HOX$
(3) $2H_2O_2 \rightarrow 2H_2O + O_2$
(4) $H_2O_2 + RH \rightarrow H_2O + ROH$

In reaction (1) H_2O_2 is reduced to water while a one-electron donor (AH_2) is oxidized to a radical (·AH). Halides (X^-) like in reaction scheme (2) are oxidized to their corresponding hypohalous acids (HOX). A few peroxidases make use of a catalatic reaction mechanism like in (3) where a hydrogen peroxide molecule reacts with another hydrogen peroxide - which acts as two-electron donor - releasing dioxygen. Reaction (4) is the peroxygenation reaction; the introduction of functionalities into organic molecules. Human peroxidasin 1 uses the reaction shown in (2) in order to catalyze the formation of sulfilimine bonds within collagen IV. Peroxidasins belong to the subfamily 2 of the peroxidase-cyclooxygenase superfamily (Soudi et al., 2012).

Human heme-containing peroxidases play an essential role in the innate immune system (MPO, EPO and LPO) as well as in hormone biosynthesis (TPO). Among prokaryotes and eukaryotes, non-heme peroxidases are widespread. Nonetheless, peroxidases containing the heme group are probably the most abundant (Zámocký et al., 2008). MPO is present in circulating neutrophils and monocytes and has antimicrobial activity. EPO is a protein of eosinophils - cells specialized for the elimination of parasites which invade tissues. TPO is a necessary enzyme in the synthesis of the hormones triiodothyronine and thyroxine catalyzing the iodination of tyrosine residues.

1.2. Human Peroxidasin 1

Peroxidasin was first discovered in hemocytes of *Drosophila melanogaster* as a constituent of basement membranes combining peroxidase and extracellular matrix motifs (Nelson et al. 1994). Peroxidasins belong to the subfamily 2 of the peroxidase-cyclooxygenase superfamily (Soudi et al., 2012). Gene analysis revealed the presence of a signal peptide, one leucine-rich repeat domain (LRR), four immunoglobulin-like motifs (Ig), a peroxidase domain (POX) and a C-terminal von Willebrand factor C domain (VWC) (Soudi et al., 2015) (Figure 1). Human peroxidasin 1 (hsPxd1) - also known as vascular peroxidase 1 (VPO) - is therefore a multidomain heme-peroxidase involved in the consolidation of basement membranes, a specialized form of the extracellular matrix (ECM). The ECM separates mesothelium and endothelium from subjacent connective tissue. Basement membranes confer structural integrity to tissues, serve as anchor point for molecules and also function as mediator for cell adhesion, growth, migration and differentiation processes. Peroxidasin, a homotrimeric iron protein with a covalently bound ferric high spin heme, uses bromide as a cofactor for the formation of sulfilimine (S=N) cross-links between two adjacent collagen IV protomers within their non-collagenous (NC) regions. This unique bond between the sulfur of methionine-93 (Met93) and the nitrogen of lysine-211 or hydroxylysine-211 (Lys211/Hyl211) amino acid residues stabilizes NC1 interfaces of contiguous triple helical collagen IV protomers. The formation of this collagen "knot" by peroxidasin requires a halide (Br^-) and releases hypohalous acid (HOBr) - a rather surprising fact since this highly reactive and usually destructive oxidant acts here in an anabolic manner (see Figure 2). Hypobromous acid reacts with the sulfur atom of Met93 thereby forming a halosulfonium cation intermediate which in a next step forms a sulfilimine bond with Lys211 or Hyl211 (Péterfi and Geiszt, 2014) (Bhave et al., 2012).

Recombinantly expressed human peroxidasin 1 (hsPxd1) is an oligomeric protein which is linked by disulfide bonds. Size Exclusion Chromatography with Multi-Angle Light Scattering (SEC-MALS) elucidated an average size of 504 kDa for this homotrimeric and glycosylated protein whereby the glycosylation pattern (11 N-glycosylation sites) was predicted computationally (Soudi et al., 2015).

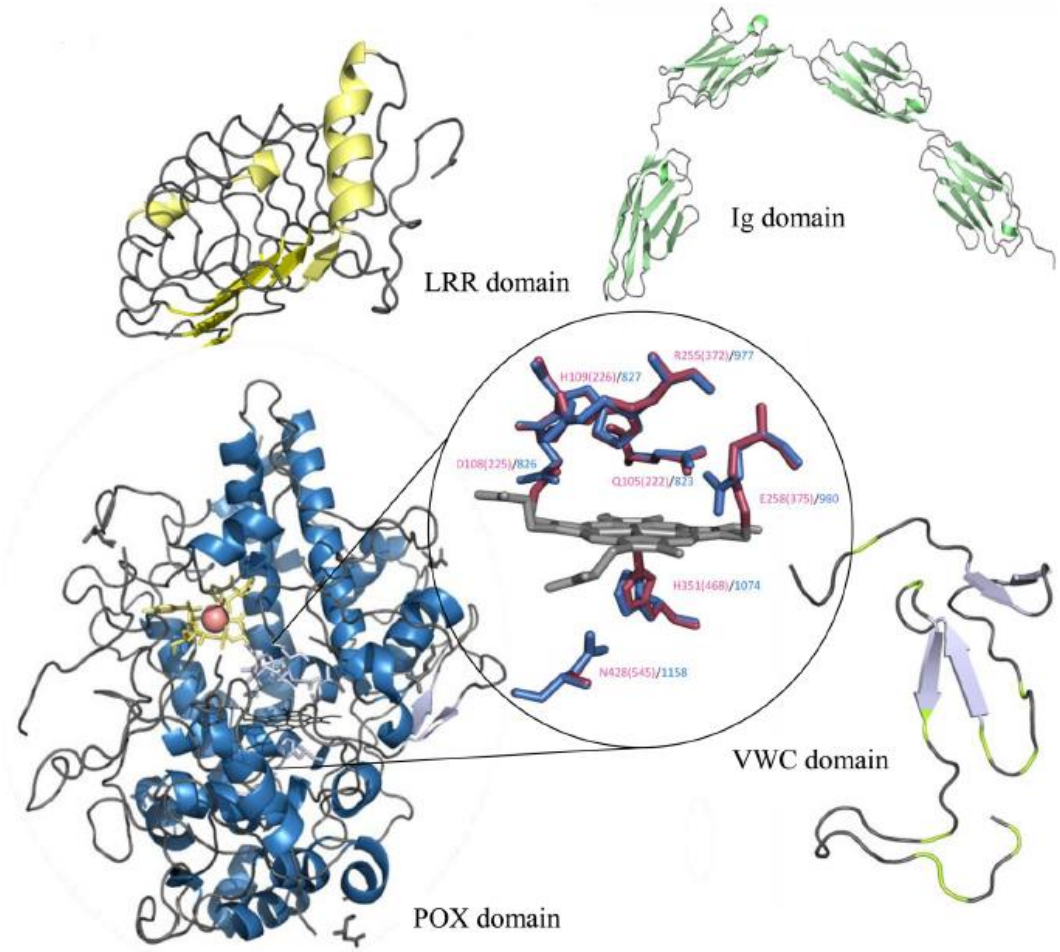


Figure 1. Models of human peroxidasin domains. The active site of the peroxidase domain is shown from gLPO (lactoperoxidase from goat, in pink) and from hsPxd (colored in blue). Figure from Soudi et al., 2015.

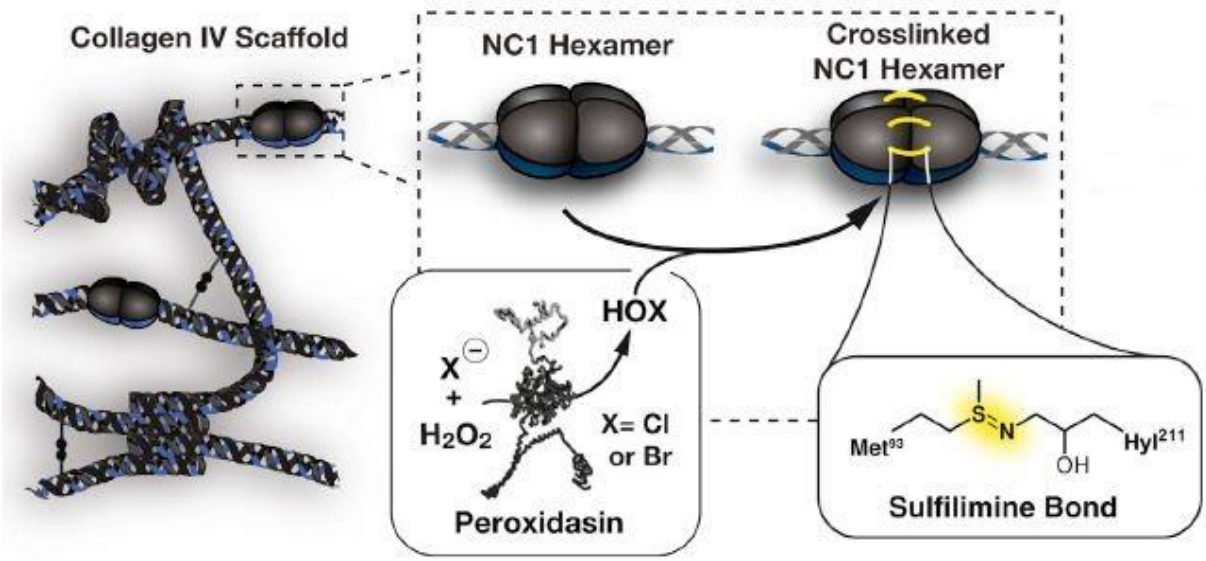


Figure 2. Overview of a collagen IV scaffold and peroxidasin mediated sulfilimine bond formation via the NC1 domain at Met93 and Lys/Hyl211. Figure from McCall et al., 2014.

1.3. Sso7d - a Potent Scaffold Protein

The basic DNA-binding protein Sso7d from the hyperthermophilic archaeon *Sulfolobus solfataricus* comprises suitable characteristics as a binding scaffold. Besides its small size of 7 kDa and a high thermal stability (T_m of 98°C) it contains no cysteines and no N-glycosylation sites. The last two properties enable bacterial expression as well as intracellular applications. High thermal stability, like in Sso7d, is an advantage for the task of protein engineering. Stable proteins are less prone to lose their native fold during the protein engineering process, because of their higher tolerance to a broader range of mutations. A triple-stranded anti-parallel beta-sheet, an orthogonal double-stranded beta-sheet (both sheets connected by short loops) and a short C-terminal alpha-helix build up the protein, which has a similar topology to that found in Src homology-3 (SH3) domains in eukaryotes. The engineered binding site on Sso7d is located on the rigid triple-stranded β-sheet on the surface of the protein whereas in antibodies and most other alternative scaffolds the binding surfaces consist of flexible loops. The rigidity of such a paratope reduces the entropic penalty upon binding and might reduce the tendency to aggregate (Baumann et al., 1994) (Traxlmayr et al., 2016).

During this study, a reduced charge variant was used (reduced charge Sso7d, rcSso7d) from a previous study (Traxlmayr et al., 2016). Since Sso7d is a DNA-binding protein it is highly positively charged (14 out of 63 residues being lysines). To achieve a charge-neutralized variant of Sso7d the two terminal lysines were deleted and four lysines were mutated to neutral amino acids while maintaining high thermal stability. Furthermore, flattening of the binding surface could also be achieved by exchanging lysine on position K39 for alanine.

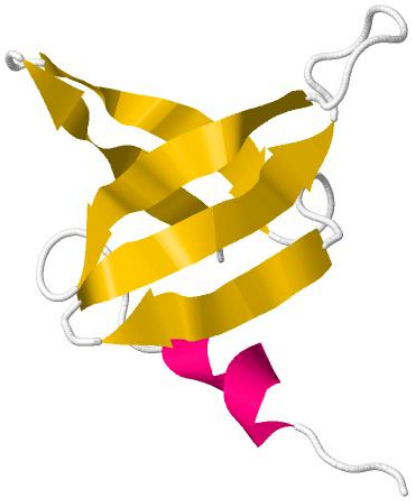


Figure 3. Cartoon model of rcSso7d (RCSB-PDB code: 1SSO). The triple-stranded anti-parallel beta-sheet in the foreground, from which nine amino acids are directed to the surface, can be engineered for binding to different epitopes (Bauman H. et al., 1994)

2. Aim of the thesis

To comprehensively characterize a protein it is necessary to elucidate its structure as well as its function. The main goal of this thesis was the development of binder proteins based on the small and thermostable scaffold-protein Sso7d for the molecular recognition and binding of human peroxidasin 1. Due to the fact that the recombinant expression of the full-length heme-containing enzyme peroxidasin 1 yields insufficient amounts of protein for detailed characterization, the main idea was to engineer small protein binders which assist in the purification process of human peroxidasin from endogenous sources, e.g. from human plasma. Furthermore, those binders can be used as crystallization chaperones and detection reagents later on.

3. Materials and Methods

3.1. Media, Buffers and Solution

In the following table all prepared reagents, buffers, media and solutions are listed.

SD-CAA medium Synthetic Dextrose CasAmino Acids	2 % w/v D-Glucose (Sigma-Aldrich®) 0.67 % w/v Difco™ yeast nitrogen base 0.5 % BD, Bacto™ Casamino Acids 70 mM citrate buffer pH 4.5
SD-CAA agar plates Synthetic Dextrose CasAmino Acids	2 % w/v D-Glucose (Sigma-Aldrich®) 0.67 % w/v BD, Difco™ Yeast Nitrogen Base, w/o amino acids 0.5 % BD, Bacto™ Casamino Acids 18.2 % sorbitol 1.5 % agar 100 mM phosphate buffer pH 6.0
SG-CAA medium Synthetic Galactose CasAmino Acids	2 % w/v D-Galactose (Carl Roth®) 0.2% w/v D-Glucose (Sigma-Aldrich®) 0.67 % w/v BD, Difco™ Yeast Nitrogen Base, w/o amino acids 0.5 % BD, Bacto™ Casamino Acids 100 mM phosphate buffer pH 6.0
YPD medium Yeast Extract Dextrose	2 % w/v D-Glucose (Sigma-Aldrich®) 1 % w/v yeast extract 2 % w/v peptone
YPD agar plates	2 % w/v D-Glucose (Sigma-Aldrich®) 1 % w/v yeast extract 2 % w/v peptone 1.5 % agar
1x PBSA buffer	1x PBS buffer 0.1 % w/v BSA (fraction V)
TEA Tris-acetate-EDTA	40 mM Tris base 20 mM acetic acid 1 mM EDTA
DTT solution (always freshly prepared) Dithiotreitol	1 M dithiothreitol
Lithium acetate	100 mM lithium acetate
SD-CAA with 30 % glycerol (cryostocks)	SD-CAA medium supplemented with glycerol to a final concentration of 30 %
LB medium Luria Broth (per litre)	2.5 g yeast extract 5.0 g peptone 5.0 g NaCl
LB agar plates (supplemented with appropriate antibiotics if necessary) (per litre)	2.5 g yeast extract 5.0 g peptone 5.0 g NaCl 7.5 g agar

YPD medium with 30 % glycerol	YPD medium supplemented with glycerol to a final concentration of 30 %)
Equilibration buffer pH 8.0	50 mM NaH_2PO_4 50 mM Na_2HPO_4 x $7H_2O$ 300 mM NaCl
1 M Imidazole Elution buffer pH 8.0 (250 mM, 15 mM and 5 mM imidazole used for elution - diluted with equilibration buffer to desired concentration)	50 mM NaH_2PO_4 50 mM Na_2HPO_4 x $7H_2O$ 300 mM NaCl 1 M imidazole
Regeneration buffer pH 5.0	20 mM MES 300 mM NaCl
SUMO storage buffer pH 7.9 (5.0/7.5/10.0 % glycerol)	20 mM Tris-Base 150 mM NaCl 3 mM β-Mercaptoethanol 1 mM EDTA 5.0/7.5/10.0 % glycerol
TB medium Terrific broth	12 g tryptone 24 g yeast extract 4 mL glycerol 2.31 g KH_2PO_4 16.43 g K_2HPO_4 x $3H_2O$ Fill up with ddH_2O to 1 L
Sonication Buffer pH 8.0	50 mM NaH_2PO_4 50 mM Na_2HPO_4 x $7H_2O$ 300 mM NaCl 3 % glycerol 1 % Triton-X100 Fill up with ddH_2O to 1 L

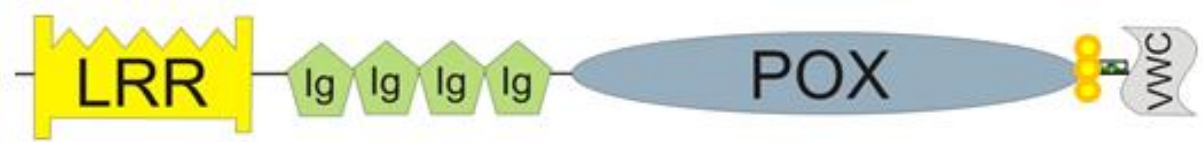
3.2. Antigen Preparation

3.2.1. Peroxidasin Constructs

Human Peroxidasin 1 (hsPxd01) is a multidomain heme protein consisting of four distinct regions: a leucine-rich-repeat (LRR) consisting of five typical repeats, four Ig domains (Ig), a heme containing peroxidase domain (POX) as well as a Von Willebrand factor C-like domain (VWC). The complete hsPxd01 is a 504-kDa protein. Because of its complexity and therefore the corresponding low expression rates in HEK293 cells (human embryonic kidney cells), unsatisfactory heme occupancy, incomplete posttranslational heme modification and low activity (Paumann-Page et al., 2017), truncated versions of the full-length human peroxidasin were expressed. Versions of the protein in possession of the linker region after the peroxidase domain form trimers. The following shortened versions of peroxidasin were used during this work (figures adapted from: Soudi et al., 2015):

Construct 1N (CON1)

This version of the protein consists of the leucine-rich repeat domain, the four Ig domains, the catalytically active peroxidase domain as well as the Von Willebrand factor C-like domain. Therefore, it is identical to the full-length version of peroxidasin (1453 amino acids).

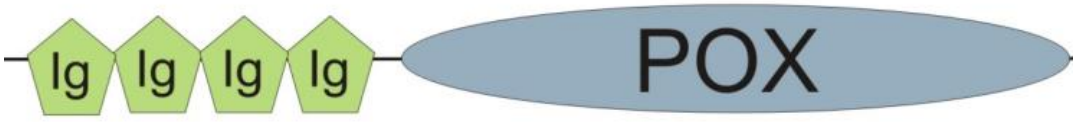


Construct 3N (CON3)

In CON3 the VWC as well as the linker region are lacking (145 kDa).

Construct 4N (CON4)

This construct consists of the peroxidase domain and the four Ig domains. It has a monomeric molecular mass of 121 kDa. This peroxidasin construct was used for all yeast display selections.



Construct 5N (CON5)

This version only contains the catalytically active peroxidase domain.

All purified peroxidasin constructs were received from Martina Paumann-Page and Benjamin Sevcnikar.

3.2.2. Biotinylation of Peroxidasin

For various yeast display selection processes, it is necessary to use a biotinylated antigen. There are two main methods available to biotinylate a target protein. The side chains of lysine residues or the free amino terminus (N-terminus) can react chemically with biotin functionalized with N-hydroxysuccinimide (NHS). The second possibility is an enzymatic linkage mediated by biotin-protein ligase (EC 6.3.4.15, BirA), which is capable of adding biotin molecules to a unique amino acid sequence (^{N}-GLNDIFEAQKIEWHE-^{C}) located on the N- or C-terminus. Care must be taken to ensure that the right buffers are used. When applying the NHC-based biotinylation method, buffers containing primary amines (like Tris or glycine) must be avoided because these components compete with the primary amines on the target protein.

In order to obtain binders for human peroxidasin, construct 4 (CON4) was used for biotinylation and subsequent yeast display selections, mainly due to its good expression rate, its high heme occupancy and its high enzymatic activity, suggesting that this construct is natively folded.

The procedure of biotinylation of the target protein (also referred to as antigen) - in this case CON4 - is necessary in order to be able to bind the antigen to streptavidin-coated magnetic beads. The high affinity between biotin and streptavidin leads to a very rigid binding of antigen to the beads which is a necessary condition in the magnetic bead selection.

The Sulfo-NHS-LC-Biotin was removed from the freezer and equilibrated to room temperature. This is important because Sulfo-NHS-LC-Biotin is moisture-sensitive. Immediately before usage, a 10 mM solution of the biotin reagent was prepared by adding 180 µL of ultrapure water to the 1 mg microtube (No-Weigh Format) and mixing by gently pipetting up and down. The calculated volume of biotin reagent was immediately added to the protein solution. The mixture was incubated for 30 minutes at room temperature.

To 2 aliquots of CON4 (2x 300 µL, 42.4 µM) a 6-fold excess (42.4 µM x 6 = 254.4 µM) of biotin was added. A biotin-stock with 10 mM was used. Therefore 7.6 µL of biotin reagent was added to 300 µL of CON4.

$$\frac{254.4\ \mu M}{10000\ \mu M} . 300\ \mu L = 7.6\ \mu L \text{ biotin reagent}$$

3.2.3. Preparative Size Exclusion Chromatography

Size exclusion chromatography (SEC) is a method used to separate molecules according to their size and shape. Smaller molecules are retained longer in the pores due to the ability to penetrate the small pores of the stationary phase (hydrophilic particles). The larger a molecule is the less it diffuses into the porous material of the stationary phase and therefore it is eluted faster from the column.

After biotinylation of CON4, three samples - the non-biotinylated construct 4 (CON4), the biotinylated construct 4 (CON4-B) and construct 5 (CON5) - were purified using a Superdex 200 HR 10/300 column with SEC buffer (20 mM phosphate, 500 mM NaCl, pH 7.44) at a flow rate of 0.4 mL/min. Prior to SEC purification all aliquots were centrifuged at 14000 rpm for 10 minutes at 4°C. The clear supernatant was applied to the column (50 µL per aliquot). The absorption was measured at 280 nm and 410 nm (heme peak of peroxidase domain).

3.2.4. Verification of CON4 Biotinylation

The HPLC was used to check the biotinylation of CON4. Therefore biotinylated CON4 (CON4-B), CON4-B in combination with streptavidin and streptavidin only was loaded on a Superdex 200 HR 10/300 with a flow rate of 0.4 mL/min (HPLC system Shimadzu LC-20AD XR). The absorption was measured at 280 nm and 650 nm.

A 3-fold molar excess of streptavidin (9.8 µL of Streptavidin-Alexa Fluor 647, 2 mg/mL, Invitrogen) was added to 15 µg (17.4 µL) of CON4-B. All samples were filtered through filter-tubes at 4000 rpm for 5 min at room temperature before 30 µL of each sample was injected onto the column.

Sample	Concentration [mg/mL]	Molecular weight [kDa]
CON4 biotinylated (CON4-B)	0.85	120.8
Streptavidin	2.0	52.8

To demonstrate that biotin was successfully coupled to CON4, streptavidin was added to the biotinylated sample. Due to the high affinity of the homo-tetrameric streptavidin molecule to biotin molecules one should see a peak shift in the chromatogram of CON4-B upon addition of streptavidin.

3.2.5. Analytic Size Exclusion Chromatography Multi-angle Static Light Scattering (SEC-MALS)

SEC-MALS couples a Multi-Angle static Light Scattering (MALS) detector with high-pressure liquid chromatography (HPLC) size exclusion chromatography (SEC). Light scattering and concentration are measured online and therefore absolute molar mass can be determined at each elution position.

SEC-MALS was used to determine if peroxidasin binders exist as monomers or tend to aggregate or stick to the column due to hydrophobic and/or electrostatic interactions and to determine their molar mass. For SEC-MALS 25 µg of protein were applied to a Superdex 200 10/300 column (HPLC system Shimadzu LC-20AD XR). The protein solutions were centrifuged before injection with Ultrafree-MC centrifugal filter devices for 3 min, 12000 g at room temperature. SEC buffer (20 mM phosphate, 500 mM NaCl, pH 7.44) was used as running buffer.

3.3. Yeast Surface Display Experiments

Yeast surface display (YSD) is a potent method to express proteins of interest on the surface of yeast and to engineer proteins used for diverse purposes like therapeutic, diagnostic, and other biotechnological applications. Biochemical and biophysical properties like affinity, specificity and stability of proteins can be altered through mutagenesis of certain regions within the protein and subsequent selection of the combinatorial libraries displayed on the yeast surface. No sub-cloning and soluble protein expression is necessary for initial characterization of enriched protein mutants. Another advantage of yeast is the ability to post-translationally modify proteins, e.g. to glycosylate them.

The protein of interest is expressed on the surface of the yeast cell as a fusion protein. The Aga2p system of *Saccharomyces cerevisiae* is most commonly used (Boder & Wittrup, 1997). In this yeast display system the protein of interest is genetically fused to Aga2p. This Aga2p protein is further linked via two disulfide bonds to the Aga1p subunit which anchors the whole complex to the cell surface. The protein of interest is encoded as an Aga2p-fusion gene and stably integrated into a single-copy yeast shuttle plasmid vector, whereas Aga1p is integrated in the chromosome. Both genes are under the control of a galactose-inducible (GAL1) promoter. Media containing galactose induce the expression of the mutant protein, which is displayed on the cell surface either linked to the C- or N-terminal end of the Aga2p subunit. Each yeast cell is capable of expressing approximately 50.000 copies of the Aga2p-fusion protein, depending of course on the properties of the fused protein. In the pCTCON2 vector, where the protein of interest is fused to the C-terminus of Aga2p, a C-terminal c-myc tag and an N-terminal hemagglutinin (HA) tag are used for immunofluorescent detection and allow for normalization of protein expression. Yeast libraries with more than 10^9 mutants can be generated using homologous recombination. Yeast cells are naturally capable of *in vivo* recircularization of a linearized vector with a corresponding overlapping insert by a mechanism called plasmid gap repair. Both – the linearized vector and the insert – are introduced into the yeast cells by electroporation. Magnetic sorting and/or flow cytometry facilitate isolation of potential binding proteins.

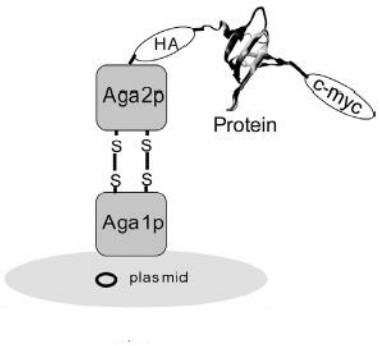


Figure 4. Schematic representation of the yeast surface display system. Two disulphide bonds link Aga2p (to which the protein of interest is coupled) to Aga1p. A c-myc as well as an HA tag allow for immunofluorescent detection by using appropriate antibodies (from: Gera, Nimish, Mahmud Hussain, and Balaji M. Rao. "Protein selection using yeast surface display."Methods 60.1 (2013): 15-26.).

3.3.1. Yeast Strain and Plasmids

The *Saccharomyces cerevisiae* yeast strain commonly used for yeast surface display is EBY100 derived from BJ5465 (*MATa, ura3-52, trp1, leu2, leu2Δ1, his3Δ200, pep4::HIS3, prb1Δ1.6R* and *can1 GAL*). EBY100 strains are Leu^- and Trp^- for selection on minimal yeast media such as SD-CAA. The Aga1p gene is stably integrated in the yeast genome. Aga1p anchors the displayed protein of interest - which is linked to Aga2p – to the cell wall. In this project the vector pCTCON2 was used. The protein scaffold - which is adjoined by two immunofluorescent tags: a hemagglutinin (HA) epitope tag at the N-terminus and a c-myc epitope tag at the C-terminus – is displayed as a C-terminal fusion to the Aga2p protein. An ampicillin resistance gene on the plasmid allows for selective growth and cloning in *E. coli* and a TRP1 gene enables selection of plasmid-positive yeast cells on minimal (i.e. tryptophane-deficient) yeast media. Besides this plasmid another vector (pCHA) may be used alternatively, if the protein of interest is linked to the N-terminus of Aga2p.

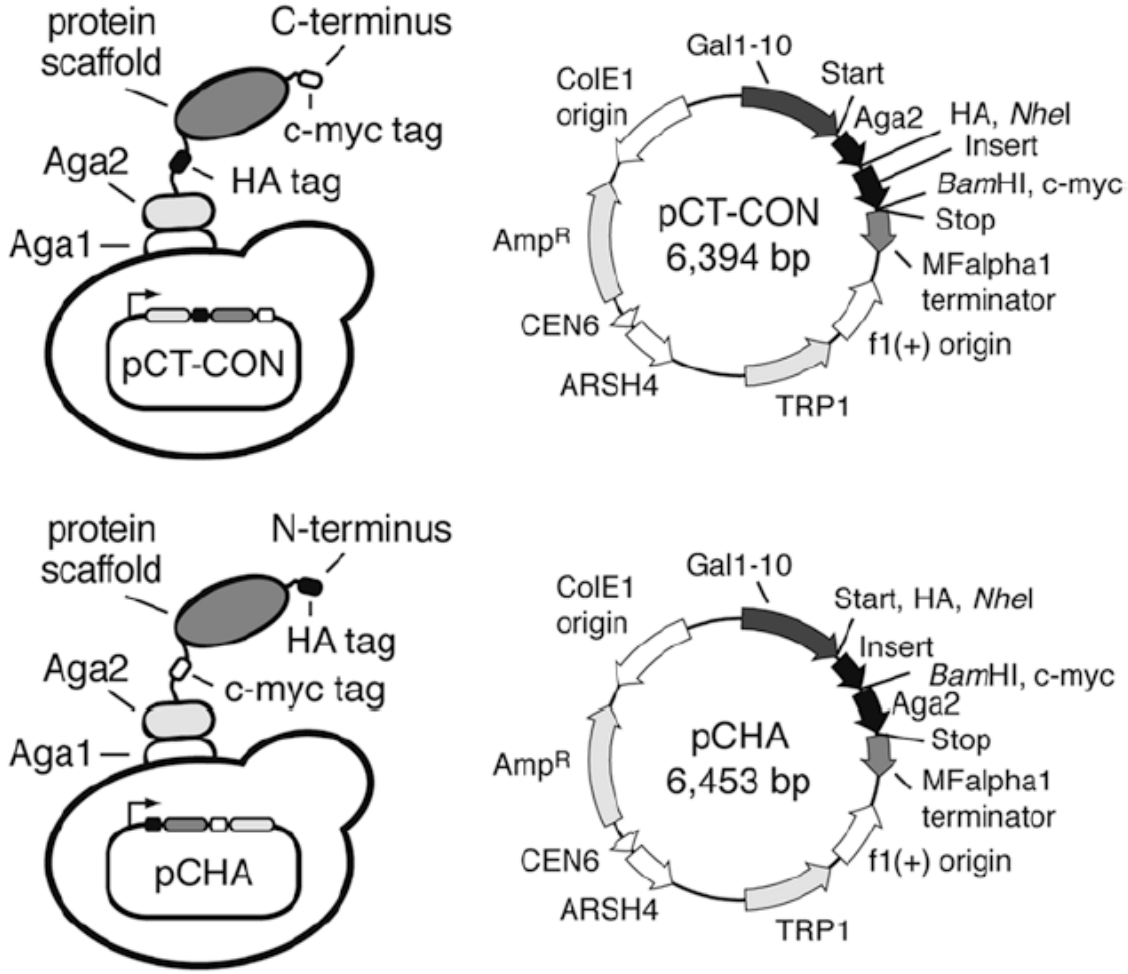


Figure 5. Graphical representation of yeast cells displaying a protein scaffold and plasmid maps of pCT-CON and pCHA (from: Angelini, Alessandro, et al. "Protein Engineering and Selection Using Yeast Surface Display." Yeast Surface Display: Methods, Protocols, and Applications (2015): 3-36.). During this work pCT-CON was used. In pCT-CON the protein scaffold contains a c-myc tag at its C-terminus and an HA-tag at the N-terminus. Aga1p anchors the displayed protein of interest - which is further linked to Aga2p – to the yeast cell surface. Both vectors contain an ampicillin resistance gene, a GAL-promoter and restriction sites for *Nhe*I and *Bam*HI.

3.3.2. Library Construction and Preparation for Selection

Generation of combinatorial libraries used in yeast surface display requires a linear DNA insert that codes for the desired protein scaffold to be mutated and engineered. Concerning stability and structure one should determine residues within the sequence which most likely tolerate mutations. Two common techniques to introduce mutations on a set of residues on the protein of interest include site-directed saturation mutagenesis with degenerated oligonucleotides and the use of nucleotide analogues in error-prone PCR. The latter has the advantage that mutation rates can be controlled by the concentration of the nucleotide analogues and the number of PCR cycles executed.

A vector e.g. pCT-CON (see Figure 5) is linearized by applying serial digestions with restriction enzymes including *NheI*, *BamHI* and *SalI*. Overlapping inserts are integrated together with the vector in the yeast cell using electroporation. Important here is that the insert contains 30 – 50 bp (or more) homology with the linearized vector at each end. The library diversity also arises from homologous recombination in yeast cells mediated by plasmid gap repair and as a result *in vivo* recircularization of the linearized vector with the overlapping inserts occurs. A further advantage of this approach is that the need for ligation and for cloning in *E.coli* is eliminated. Library sizes of more than 10^9 can be achieved.

The libraries used during this work were constructed based on the DNA-binding protein Sso7d from the hyperthermophilic archaeon *Sulfolobus solfataricus.* Due to its small size (7 kDa), its high thermal stability (T_m = 98°C) and the absence of glycosylation sites and cysteines it is a suitable binding scaffold for a wide spectrum of targets and applications. Sso7d is a highly positively charged protein with 14 out of 63 residues being lysines (22 %). To bypass non-specific binding interactions (e.g. with negatively charged mammalian cell membranes) a reduced charge variant was constructed (reduced charge Sso7d, rcSso7d) in a previous study. Starting from the charge-reduced scaffold, two libraries were constructed: rcSso7d-11 and rcSso7d-18 with each library containing $1.4x10^9$ independent transformants (Traxlmayr et al., 2016).

For the present project, both libraries (rcSso7d-11 and rcSso7d-18) were used as starting libraries for the construction of binders for molecular recognition of human peroxidasin 1. To cover at least 5x the library diversity, six aliquots á 1 mL per library were thawed. The cells were transferred into a 50 mL Falcon tube filled with 30 mL SD-CAA. Every tube with cells was washed again with 1 mL SD-CAA to get as many cells as possible. The final volume per library was 42 mL. For both libraries 100 µL for the 10^{-5}- and 100 µL for the 10^{-6}-dilutions were plated on SD-CAA agar plates. The plates were incubated at 30°C for 3 days. For each library 2 2-litre flasks containing 1 L of SD-CAA were incubated with 21 mL of the above mentioned cell suspension overnight at 30°C and 180 rpm. On the next day optical density at 600 nm (OD_{600}) was measured for all 4 flasks. At least tenfold the estimated library diversity ($1.4x10^{10}$ cells from each library) were passaged into a total of 1 L SD-CAA (1 flask per library). Glycerol stocks à 1.5 mL containing 15% glycerol were prepared of both libraries as a backup and immediately stored at -80°C.

Library rcSso7d-11 and library rcSso7d-18 were incubated until an OD_{600} of approximately 4 was reached (mid-log phase, 3 – 5 hours). An OD_{600} of 1 corresponds to approximately 10^7 cells per mL. $1.4x10^{10}$ cells (tenfold diversity) were centrifuged in 50 mL Falcon tubes (1500 g, 5 min). The supernatant was discarded and the cell pellet was resuspended in 1 L SG-CAA per library for overnight induction at 20°C, shaking at 180 rpm (medium contains galactose as carbon source which causes the GAL1 promoter to become activated and therefore inducing transcription of the gene of interest). Cells can be used either directly for selection or stored at 4°C for up to 1 month.

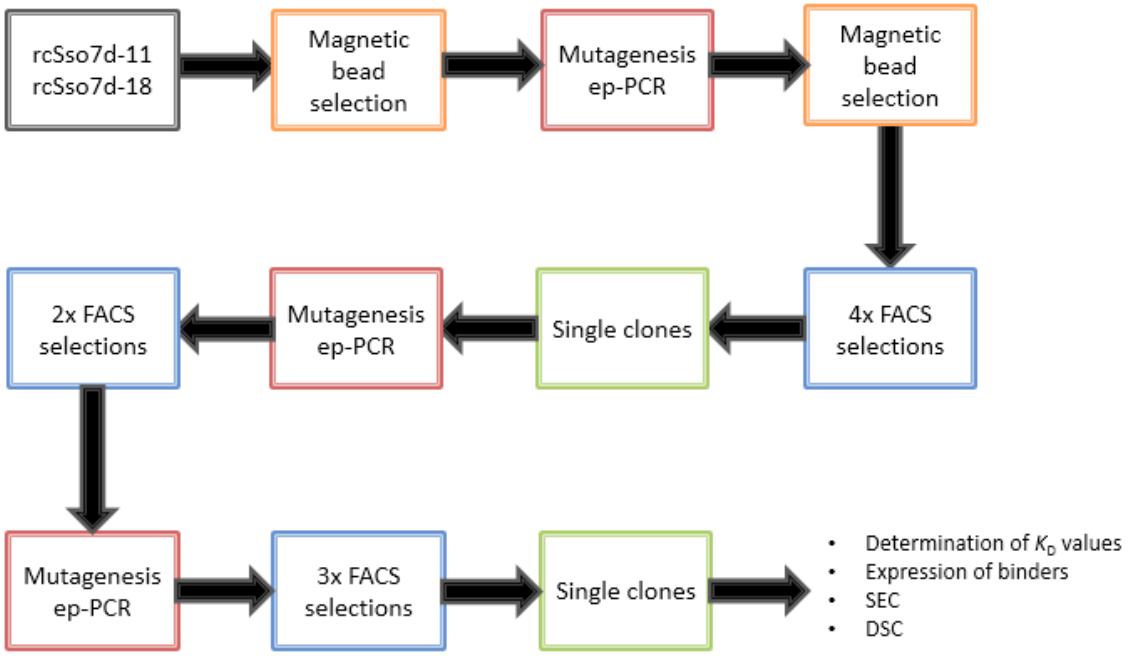


Figure 6. Overview of the whole selection process using two yeast libraries (rcSso7d-11, rcSso7d-18) as starting point. The two libraries were combined/pooled prior to the first magnetic bead selection. The next steps included the mutagenesis by ep-PCR (error-prone PCR), several FACS (fluorescent activated cell sorting) selection rounds and the isolation of single clones as well as the determination of characteristics of these single clones including K_D values as well as biophysical parameters using DSC (differential scanning calorimetry) and SEC (size exclusion chromatography). In the end, we were able to select 11 binders subdivided into 4 families with varying strength in affinity for their antigen.

3.3.3. Magnetic Bead Selection

To select protein binders from a naïve yeast display library a two-step procedure is performed: first part of the process is an (antigen-coated) magnetic bead selection where weak protein binders are isolated in a short period of time at high throughput. The magnetic beads are 2.8 µm superparamagnetic beads with the ability to bind biotinylated ligands. After the magnetic bead selection cells are screened using flow cytometry sorting.

The magnetic bead selection is divided into two steps: a negative and a positive selection. During the negative selection, protein binders are incubated with bare streptavidin-coated magnetic beads (without biotinylated antigen). In this step one can deplete binders against streptavidin, as well as binders that bind non-specifically to the beads. In the positive selection the biotinylated antigen is immobilized on the magnetic beads and incubated with the (previously depleted) yeast library. The yeast-bead complexes are isolated using a magnet in the end.

Due to the fact that about $5x10^4$ proteins are displayed on the surface of the yeast cell which can interact with multiple copies of biotinylated antigen coupled to streptavidin-coated micron-sized magnetic beads, it is possible to isolate weak binders because of this highly avid interaction between yeast cells and magnetic beads.

On the first selection day only one round of positive selection (with antigen-loaded beads) was performed, but no negative selection with bare beads. All magnetic bead selections were performed in 1.5 ml microcentrifuge tubes. Two times 100 µL of magnetic beads (Dynabeads® Biotin Binder) were washed with 900 µL PBSA, inverted several times and placed on a magnetic rack (DynaMag™-2 Magnet) for 2 minutes. After the 2 minutes, the supernatant was pipetted off. After a second washing step with 1000 µL buffer the beads were resuspended in a total of 1000 µL buffer containing 330 pmoles of biotinylated antigen (44.8 µL CON4-B + 953.2 µL PBSA). The magnetic beads and biotinylated CON4 were incubated in a 1.5 ml microcentrifuge tube for > 2 hours on a rotating wheel.

During this incubation time cell density of both libraries (incubated for > 16 h in SG-CAA) was measured. From both libraries $1.2x10^{10}$ cells were centrifuged at 2000 g for 10 minutes at 4°C. The cell pellet was resuspended in 20 mL PBSA per library. From this volume 1 mL of a 1:20 dilution was stored at 4°C for later FACS analysis. Both libraries were centrifuged at 1500 g for 5 minutes at 4°C and subsequently resuspended in 10 mL of PBSA per library. Now both libraries rcSso7d-11 and rcSso7d-18 were pooled into one falcon tube (20 mL). This cell suspension was split into 20 equal parts (á 1 mL) in 1.5 ml microcentrifuge tubes ($1.2x10^9$ cells per tube). Cells were centrifuged at 2000 g and 4°C for 3 minutes. All cell pellets were resuspended with 800 µL buffer to reach a final volume of 950 µL per tube.

After 2 hours of incubation on a rotating wheel the antigen-coated beads were placed on the magnetic rack for 2 minutes, buffer was removed and the magnetic beads washed again with 1000 µL buffer two times. The antigen-beads were resuspended in 550 µL of buffer. 50 µL of beads were added to each tube of the pooled cells from libraries rcSso7d-11 and rcSso7d-18. After an incubation time of > 2 hours, yeast cells and antigen-

beads were placed on the magnetic rack for 2 minutes, after which the unbound cells were discarded. After a final washing step in 1000 µl PBSA the cell-bead mixture was resuspended in 1 mL SD-CAA per tube. The library was incubated in a total volume of 140 mL SD-CAA at 30°C shaking at 180 rpm overnight (> 16 hours). 50 µL of a 1:20 and a 1:200 dilution were plated on SD-CAA agar plates.

To avoid carry-over of antigen-loaded magnetic beads to the following negative bead selection round, antigen-loaded beads were separated from the selected and expanded yeast cells on the day after the first magnetic selection. For that purpose, 10^9 cells were centrifuged (2000 g, 3 min, RT), resuspended in 2 mL SD-CAA and placed on the magnetic rack for 2 minutes. The unbound cells were taken off, pipetted into a new micro-centrifuge tube and the separation process using the magnet was repeated. Subsequently, the unbound cells were diluted in 70 mL SD-CAA and incubated at 30°C until an OD_{600} of 3 - 4 was reached. Once more 10^9 cells were centrifuged (1500 g, 5 min, RT), resuspended in 100 mL SG-CAA for induction and incubated overnight (20°C, 180 rpm). The remaining volume of the cells was further incubated in SD-CAA.

From the induced library P0.1 (P stands for the name of the library "peroxidasin binders". The number 0.1 means that the library was zero times mutagenized using error-prone PCR and one time selected using the magnetic bead selection. In later rounds of selection FACS was used in order to sort out high affinity binders.) $2x10^9$ cells were centrifuged (1500 g, 5 min, 4°C), resuspended in 2 mL PBSA and split into two 1.5 mL tubes (1 mL each). Subsequent to an additional centrifugation (2000 g, 3 min, 4°C) the cell pellet (ca. 100 µL) was resuspended in 850 µL PBSA to reach a final of volume of 950 µL. A sample of a 1:20 dilution was prepared in parallel for FACS analysis.

Next, 30 µL of Biotin Binder Dynabeads were mixed together with 970 µL PBSA, placed on the magnetic rack, washed 2 times with PBSA and resuspended in 170 µL of buffer per tube. Now 3 rounds of negative selection using the bare beads were executed. Therefore 50 µL of washed bare beads were added to one of the two tubes containing the PBSA-washed yeast cells and each time incubated for 1.5 hours. After each incubation step, the unbound cells were transferred to a fresh tube together with 50 µL of uncoated beads.

Following these 3 rounds of negative selection the unbound cells were incubated with antigen-coupled beads at 4°C for > 2 hours on a rotating wheel.

The second tube of library P0.1 was selected a second time against biotinylated antigen (not part of the 3 times negative selection). This time the magnetic beads were coupled with 66 pmoles of biotinylated CON4 in 200 µL PBSA. After >2 hours incubation on a rotating wheel, 800 µL PBSA were added to the bead-antigen mixture, placed on the magnet for 2 minutes, washed 2 times with 1 mL PBSA and resuspended in 120 µL buffer. Subsequently, 50 µl of the antigen-loaded beads were added to the cells, followed by incubation at 4°C on a rotating wheel for 2 hours. Tubes were placed on the magnet and unbound cells discarded. Cells were now washed with 1 mL PBSA, mixed and placed on the magnet in order to get rid of unbound cells. After repeating this process 2 times the

beads and cells were incubated in 30 mL SD-CAA (2 libraries: P0.2 and P0.2 depl.[1]) overnight at 30°C on an orbital shaker (180 rpm). Dilutions for library diversity determination (undiluted and 1:10) were plated on SD-CAA plates and incubated at 30°C.

After the first round of mutagenesis by error-prone PCR and transformation into yeast (libraries P1.0 and P1.0 depl.) another round of magnetic bead selection was performed as described before. Again, one library was additionally selected against bare beads (depleted library) whereas the other library (P1.0) was only selected against beads incubated with 66 pmoles of biotinylated CON4.

3.3.4. Plasmid Extraction from Yeast Cells

The extraction of plasmids from yeast cells was performed with the Zymoprep™ Yeast Plasmid Miniprep II kit. The extracted and purified DNA can be used for sequencing purposes to identify the amino-acid sequence of potential binders (after transformation of *E. coli*), or as a template to prepare mutagenized DNA for further library generation.

0.5 – 1.0 x 10^8 yeast cells from a growing culture with an $OD_{600} \sim 2 – 3$ were aliquoted into 1.5 mL microcentrifuge tubes. Cells were spun down at 12000 rpm for 1 minute. Supernatant was discarded and 200 µL of Solution 1 was added to the pellet (pipetting up and down gently several times). 6 µL of Zymolyase enzyme was added to the tube and gently mixed. Cells were incubated at 37°C for 3 – 5 hours and tubes mixed around every 20 minutes. Afterwards 200 µL of Solution 2 and 400 µL of Solution 3 were added and mixed by pipetting up and down gently. Cell debris was centrifuged at 12000 rpm for 10 minutes at room temperature. The supernatant was transferred into a new 1.5 ml microcentrifuge tube and centrifuged again at 12000 rpm for 3 – 5 min depending on the amount of debris remaining. The clear supernatant was transferred to a silica spin-column and centrifuged at 12000 rpm for 1 minute. Flow through was discarded and the column was washed with 550 µL of wash buffer and spun for 2 minutes. The column was placed into a new 1.5 mL tube and 10 – 15 µL of HQ-H_2O was added, incubated for 1 minute at room temperature and immediately centrifuged for 1 minute at 12000 rpm. Plasmids were stored at -20°C.

[1] P0.2 depl. (depleted) were those cells which were three times negatively selected against bare beads.

3.3.5. Mutagenesis by Error-prone PCR of DNA Encoding Peroxidasin Binders

In total 3 rounds of mutagenesis by error prone PCR were executed under the conditions described in Table 1 and Table 2. The first round of mutagenesis was executed after the first two rounds of magnetic bead selection. The second round of mutagenesis was done based on 4 different single clones after the 4th FACS selection and the third one after the 6th FACS round.
DNA primers used for the mutagenesis of the insert DNA were the same as the ones used for amplification and generation of the DNA library via homologous recombination in the vector pCTCON2 (30 - 50 base pairs homology region at each end of the linearized vector with the primers). Either 10 ng from miniprep or 5 µL of Zymoprep served as DNA template.
As a negative control ddH_2O was used instead of a DNA template. As a positive control 1 µL of plasmid E11.8 pCTCON2 was used.

epSso_fwd:

5′-GGCTCTGGTGGAGGCGGTAGCGGAGGCGGAGGGTCGGCTAGC -3′

epSso_rev:

5′-CTATTACAAGTCCTCTTCAGAAATAAGCTTTTGTTCGGATCC -3′

Table 1. Components for error prone PCR reaction mix.

Component	Initial concentration	Final concentration	Volume [µL]
epSso_fwd	10 µM	0.5 µM	2.5
epSso_rev	10 µM	0.5 µM	2.5
dNTPs mix	10 mM	200 µM each	1.0
8-oxo-dGRP	20 µM	2 µM	1.0
dPTP	20 µM	2 µM	1.0
ThermoPol reaction buffer	10x	1x	5.0
DNA template	1 ng/µL	-	-
Taq DNA Polymerase	5 U/µL	0.05 U/µL	0.5
ddH_2O	-	-	Fill up to 50 µL
Total volume	-	-	50 µL

Table 2. Cycling conditions for error prone PCR.

Segment	Cycles	Temperature [°C]	Time [sec]
1	1	94	180
2	15	94	45
		60	30
		72	60
3	1	72	600

3.3.6. Amplification of Mutagenized DNA by PCR

After purifying the DNA from the agarose gel the insert was amplified by PCR under the conditions given in Tables 3 and 4.

Table 3. Components for amplification by PCR of mutagenized DNA.

Component	Initial concentration	Final Concentration	Volume [µL]
epSso_fwd	10 µM	1 µM	5.0
epSso_rev	10 µM	1 µM	5.0
dNTPs mix	10 mM	200 µM each	1.0
ThermoPol reaction buffer	10x	1x	5.0
DNA template (from gel)	-	-	5.0
Taq DNA Polymerase	5 U/µL	0.05 U/µL	0.5
ddH_2O	-	-	Fill up to 50 µL
Total volume	-	-	50 µL

Table 4. Conditions for amplification PCR.

Segment	Cycles	Temperature [°C]	Time [sec]
1	1	94	180
2	30	94	45
		60	30
		72	60
3	1	72	600

3.3.7. Agarose Gel Electrophoresis of PCR Products

In order to confirm the right length of PCR products they were analyzed by agarose gel electrophoresis. The principle behind agarose gel electrophoresis is based on the migration of negatively charged DNA towards a positive electrode when an electric field is applied. How fast a DNA molecule migrates towards the positive electrode depends on its size, the degree of polymerization of the gel and the applied voltage.

The gel is composed of TAE (Tris-acetate-EDTA) buffer, 1% (analytical gel) or 2% (preparative gel) agarose and SYBR® Safe DNA gel stain (Life Technologies). 5 or 10 µL of PCR product as well as a 1:5 diluted DNA ladder were mixed with 1 or 2 µL of DNA loading dye (both from Fermentas) and applied to the wells of the gel. The electrophoresis was run approximately 45 min at 120 V. The DNA bands on the gel were visualized using a GeneFlash with UV transilluminator from Syngene.

3.3.8. Purification of PCR-products

The PCR products were precipitated using ethanol. To achieve this, the pH of the DNA-sample was adjusted by adding 10% 3 M sodium acetate (pH 5.2) of the sample volume. Then at least 2x the sample volume ethanol (100 %) was added and incubated for 2 minutes at room temperature (alternative incubation: at -20°C overnight or a mixture e.g. 15 min at -20°C and 5 min at RT). Samples were centrifuged at 14000 rpm for 5 min, the supernatant was removed and 500 µL of 70 % ethanol added and briefly mixed. After another centrifugation step and removal of the supernatant 500 µL of 100 % ethanol was added, vortexed briefly and centrifuged under the conditions mentioned above. The supernatant was removed and the pellets air dried for some hours until no liquid was left. The dried pellet was dissolved in 10 µL of ddH_2O.

3.3.9. Transformation of *E.coli* and *S. cerevisiae*

Yeast Transformation I (Electroporation)

Ethanol precipitated inserts which were mutagenized by error prone PCR were used to transform EBY100 yeast cells using electroporation for further rounds of selection. For that purpose, 150 mL of YPD medium were inoculated at 30°C shaking with 180 rpm with an EBY100 overnight culture at an OD_{600} = 0.2 (50 mL are sufficient for two electroporations). As the culture reached an OD_{600} = 1.3 – 1.5 the cells were pelleted at 1500 g for 3 minutes (50 ml per tube). After removal of the supernatant the cells were resuspended in 25 mL 100 mM lithium acetate per tube. Freshly prepared dithiothreitol (DTT) was added to the cells to a final concentration of 10 mM and incubated at 30°C while shaking for 10 minutes. Cells were pelleted at 1500 g for 3 minutes, supernatant removed and cells placed on ice. For the following steps it was important to work on ice and to use chilled reagents. The EBY100 cells were resuspended in 25 mL sterile H_2O per tube and again centrifuged with 1500 g for 3 minutes. The pellet was resuspended in 250 µL of chilled sterile H_2O per tube. The total volume was approximately 500 µL of cells per tube, i.e. for each 50 mL of original culture volume. In parallel 4 µg of vector pCT-CON2 (linearized by *Sal*I, *Nhe*I, *Bam*HI) was added to amplified inserts. To ensure that DNA is completely dissolved (from ethanol precipitation) 250 µL of cells were mixed with the DNA inserts. The mixture was transferred to pre-chilled electroporation cuvettes and was kept on ice till electroporation. A single pulse at 500 V with a 15 ms pulse duration was executed using the square wave protocol on the Bio-Rad Gene Pulser XCell. Cells were immediately rescued with 1 mL of YPD medium and transferred to a 15 mL tube. The cuvette was rinsed with additional 1 mL of YPD and transferred to the same tube. After 1 hour of incubation at 30°C without shaking cells were plated and grown on SD-CAA at 30°C for 2 – 3 days in order to determine the number of transformants of the new library. Usually this protocol yields 10^7 – 10^8 transformants. The remaining cells were pelleted at 1500 g for 3 minutes and resuspended in 100 mL SD-CAA (incubated at 30°C, 180 rpm overnight).

Yeast Transformation II (Frozen-EZ Yeast Transformation II™)

To transform *Saccharomyces cerevisiae* (EBY100) with purified plasmids of individual clones, the Frozen-EZ Yeast Transformation II™ kit was used. With this kit competent cells can be prepared rapidly. EBY100 cells were grown at 30°C in 10 mL YPD until an OD_{600} of 0.8 - 1.0 was reached. Cells were pelleted at 500 g for 4 minutes and the supernatant was discarded. 10 mL of EZ1 solution was added to wash the pellet, which was repelleted after this washing step. 1 mL of EZ2 solution was added to resuspend the pellet. The cells could be used immediately or stored at -80°C.

To perform the transformation 10 µL of competent cells were mixed with 1 - 2 µL of plasmid DNA (0.2 – 1.0 µg). After addition of 100 µL of EZ3 solution and thorough mixing the cells were incubated for 45 – 60 min at 30°C. 80 – 100 µL were spread on SD-CAA plates and incubated at 30°C for 2 – 4 days.

Transformation of *E.coli* cells

For the purpose of transformation 50 µL of heat shock competent cells (CL-10CK) were mixed with 2.5 µL of plasmids isolated from yeast cells. As negative control only cells were used without insert. The cells were incubated on ice for 20 minutes and heat shocked for exactly 30 seconds at 42°C in a prewarmed waterbath. Cells were placed on ice for 2 minutes and immediately 500 µL of prewarmed (37°C) SOC medium was added to the tubes. After incubation at 37°C for 40 min aliquots of 50 – 100 µL were plated on LB agar containing ampicillin (100 µg/mL). The plates were incubated at 37 °C overnight.

Sequencing of Mutated POX-Binders

All sequencing steps were carried out by Microsynth AG using Sanger sequencing technology. Single colonies were picked from LB plates containing 100 µg/mL ampicillin and inoculated in 5 mL LB medium containing 100 µg/mL ampicillin overnight at 37°C, 180 rpm. The plasmids were isolated using the Illustra™ plasmidPrep Mini Spin Kit from GE Healthcare according to the manufacturer's protocol except for the elution step (HQ-H_2O was used instead of elution buffer). The following forward-primer (10 µM) was used for sequencing:

CON2_seq_MWT 5'-CGTTTGTCAGTAATTGCGGTTCTC-3'

3.3.10. Fluorescence Activated Cell Sorting (FACS) of Libraries

A fluorescence-activated cell sorter has the ability to analyze or separate cells in suspension based on the different fluorescent dyes applied and the morphology of the cells. FACS machines possess the capability to analyze and sort up to 10^8 cells per hour. Cells within a fluid stream are led in single file through a nozzle where they immediately pass a laser beam and the emitted fluorescence is detected by a photocell. Due to vibration the stream of cells breaks into droplets. An electric charge is applied on cells of interest detected by the system allowing separation through a high-voltage deflection system. Cells of interest are collected in a sample tube and used for further applications.

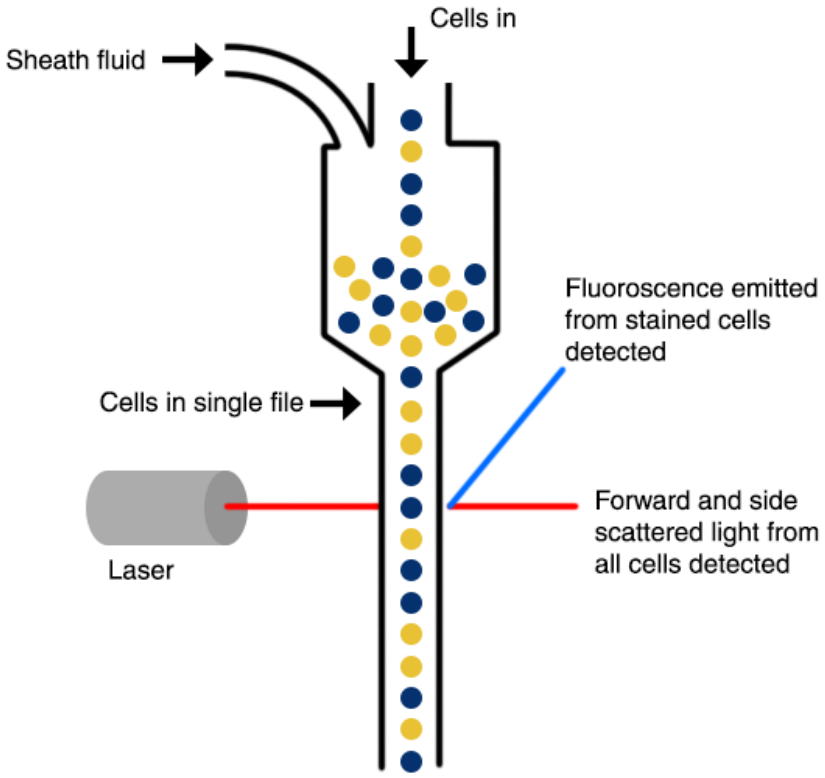


Figure 7. Schematic overview of a fluorescent activated cell sorter.
Figure from: https://www.abbexa.com/flow-cytometry-info

After the isolation of yeast cells using the magnetic bead selection cells are further screened and isolated using FACS. Flow cytometry sorting is performed to select yeast cell clones, which are displaying full-length binders in a more stringent way compared to magnetic bead selection. In subsequent FACS rounds binders are selected with higher affinity for the antigen (peroxidasin).

The yeast display vector used (pCTCON2, 6394 bp) contains two epitope tags at the N-terminus (hemagglutinin, HA) and C-terminus (c-myc) of rcSso7d, respectively, which allows for detection of full-length binders (c-myc) as well as total expression level (HA). Yeast clones displaying a full-length binder result in a dual positive signal for both tags whereas yeast clones displaying truncated protein scaffolds (e.g. due to nonsense or frame shift mutations) will result in a single positive fluorescence signal.
Full-length protein binders are then screened for protein expression level and binding to the target protein (peroxidasin and its constructs). Yeast cells displaying full-length protein binders which concomitantly bind to the antigen will result in a double positive signal for the tag and the target. By applying this two-color labeling the binding signal for the target protein can be normalized to cell surface expression levels and allows quantitative selection in real time. Since a correlation between protein surface expression and protein stability exists, it is possible to select for both parameters (i.e. affinity and stability) contemporaneously.

For preparation of yeast cells for flow cytometric sorting, all steps were performed using ice-cold PBSA buffer and a pre-cooled table top centrifuge. All tubes must be protected from light as soon as fluorescent reagents have been added.
Depending on the library size an appropriate number of cells was obtained from overnight yeast cultures induced in SG-CAA at 20°C (at least tenfold of the estimated library diversity) and subsequently centrifuged with 2000 g for 3 minutes at 4°C. The cells were mixed with 300 µL PBSA and centrifuged with 2000 g for 3 minutes at 4°C. After a washing step with 1 mL PBSA the cells were resuspended in PBSA containing the corresponding primary detection reagents and incubated at 4°C while rotating for 60 minutes. After centrifugation and washing steps the cells were resuspended in PBSA containing secondary detection reagents and incubated for 30 minutes at 4°C on a rotating wheel. After a final washing step with 1 mL of PBSA cells were centrifuged and resuspended just before sorting. Dual positive cells (fluorescence for expression and binding) were collected in 5 mL tubes containing 2 mL of fresh SD-CAA and were recovered, passaged and induced for a subsequent selection process by flow cytometry.

In Table 5 all performed FACS selection rounds are listed with corresponding amounts of antigen used, how many cells from which library and which fluorescently labelled secondary reagents were used.

Table 5. Overview of all sort rounds using FACS. Stock concentrations of used primary and secondary labelling reagents: 9E10 anti-c-myc (0.5 mg/mL), Streptavidin-AF647 (2.0 mg/mL), Goat anti-mIgG-AF488 (2.0 mg/mL), Penta His-AF488 (0.2 mg/mL) and Anti-HA-AF647 (0.5 mg/mL).

FACS round	Libraries sorted	Number of sorted cells per library	Antigen	Concentration of antigen [nM]	Primaries dilution	Secondaries dilution
1	P1.1 P1.1 depl.	$1.1x10^5$ $1.7x10^5$	CON4-B	200	9E10 anti-c-myc 1:100 CON4-B	Streptavidin-AF647 1:200 Goat anti-mIgG-AF488 1:100
2	P1.2 P1.2 depl.	$1.3x10^4$ $1.5x10^4$	CON4-B	100	9E10 anti-c-myc 1:100 CON4-B	Streptavidin-AF647 1:200 Goat anti-mIgG-AF488 1:100
3	P1.3 P1.3 depl.	$6.0x10^3$ $6.0x10^3$	CON4	100, 10	CON4	Penta His-AF488 1:40 Anti-HA-AF647 1:100
4	P1.3 P1.3 depl. P1.4 P1.4 depl.	Not sorted Not sorted $1.1x10^4$ $3.5x10^3$	CON4 CON5	100	CON4 CON5	Penta His-AF488 1:40 Anti-HA-AF647 1:100
4a[2]	P1.4 CON5 P1.5 depl. CON5	Not sorted	CON5	100	CON5	Penta His-AF488 1:40 Anti-HA-AF647 1:100
5[3]	P2.0.1 P2.0.6 P2.0.8 P2.0.9	$1.2x10^5$ $1.2x10^5$ $1.5x10^5$ $1.5x10^5$	CON4-B	50	9E10 anti-c-myc 1:200 CON4-B	Streptavidin-AF647 1:200 Goat anti-mIgG-AF488 1:200
6	P2.1.1 P2.1.6 P2.1.8 P2.1.9	$1.7x10^4$ $3.8x10^4$ $5.1x10^4$ $2.5x10^4$	CON4-B	20	9E10 anti-c-myc 1:400 CON4-B	Streptavidin-AF647 1:200 Goat anti-mIgG-AF488 1:200
7	P3.0.1 P3.0.6 P3.0.8 P3.0.9	$6.0x10^4$ $3.2x10^4$ $6.1x10^4$ $3.9x10^4$	CON4-B	20	9E10 anti-c-myc 1:200 CON4-B	Streptavidin-AF647 1:200 Goat anti-mIgG-AF488 1:200
8	P3.1.1 P3.1.6 P3.1.8 P3.1.9	Not sorted $6.5x10^3$ $7.5x10^3$ $1.2x10^3$	CON4	4	CON4	Penta His-AF488 1:40 Anti-HA-AF647 1:100
9	P3.2.6 P3.2.8 P3.2.9	$7.5x10^3$ $4.0x10^3$ $8.0x10^3$	CON4	2	CON4	Penta His-AF488 1:60 Anti-HA-AF647 1:150

Two different staining procedures were performed depending on whether or not the antigen was biotinylated (only CON4 was used in a biotinylated version, abbreviated with CON4-B).

[2] In this FACS round no binders were selected due to the fact that none of them showed any binding to CON5.
[3] The libraries in this round are descendants from single clones which were mutagenized by ep-PCR before the FACS sort.

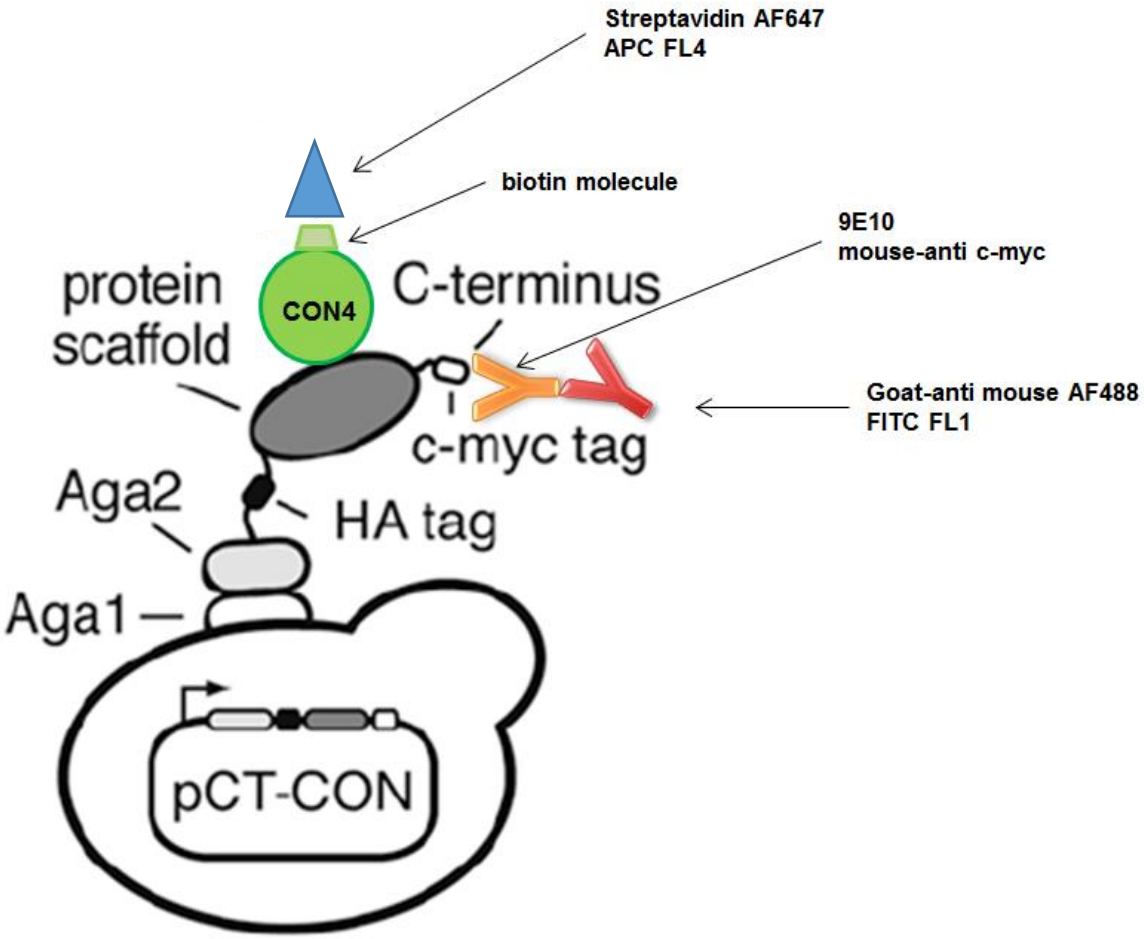


Figure 8. Staining scheme using biotinylated CON4. The c-myc tag at the C-terminus was labeled with 9E10 mouse-anti c-myc antibody and further with goat-anti-mouse-IgG Alexa-Fluor 488 as secondary detection antibody which shows cell surface expression of the potential binder. Biotinylated CON4 was added in the primary staining step and streptavidin Alexa-Fluor 647 was added to detect binding to the target protein (CON4). Figure was adapted from: Angelini, Alessandro, et al. "Protein Engineering and Selection Using Yeast Surface Display." Yeast Surface Display: Methods, Protocols, and Applications (2015): 3-36.

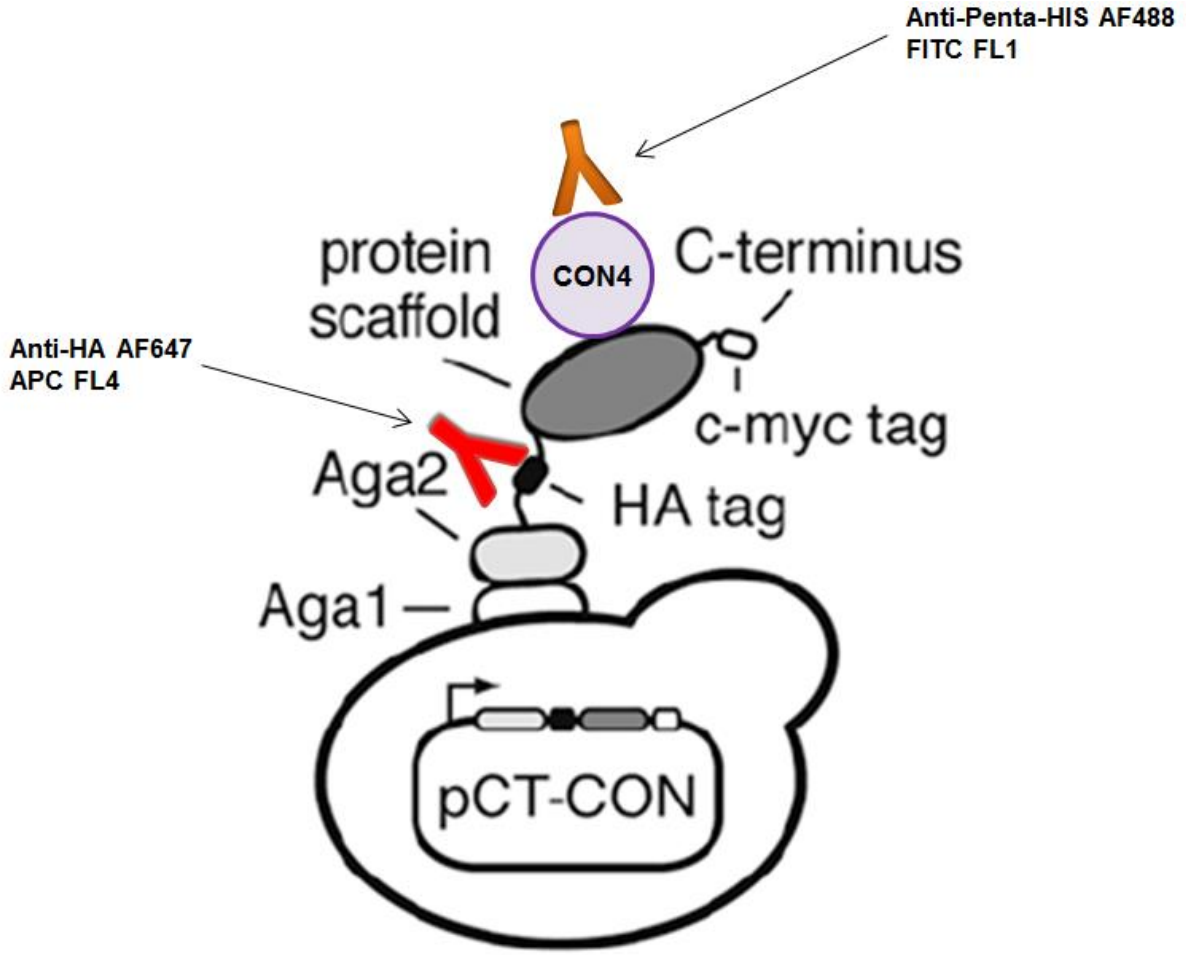


Figure 9. Staining scheme using non-biotinylated CON4 (same is valid for the staining of full-length peroxidasin, CON3 and CON5). Non-biotinylated CON4 was the only primary reagent. An anti-HA Alexa-Fluor 647 antibody was applied as secondary reagent for detection of cell surface expression level. In order to determine the interaction of binder with the target protein an anti-Penta-His Alexa-Fluor 488 antibody was utilized. This was possible because all expressed peroxidasin constructs contained a His-tag for detection. Figure was adapted from: Angelini, Alessandro, et al. "Protein Engineering and Selection Using Yeast Surface Display." Yeast Surface Display: Methods, Protocols, and Applications (2015): 3-36.

3.3.11. Binding analysis on the yeast surface

Selected protein binders for CON4 were further characterized using yeast surface titration. With this method, one can determine the K_D of the corresponding mutant protein. Yeast cells are induced and therefore express the binder on their surface. The cells are then incubated with varying concentrations of the target (CON4). The fraction of cell surface fusions bound to the antigen is determined using flow cytometry. Obtained data are fitted to a single step binding isotherm to estimate the K_D. Two important points should be mentioned. First, determination via yeast surface titration is equal to the K_D measurement using soluble protein (Traxlmayr et al., 2016). Second, even small differences in K_D can be discriminated which leads to a better understanding of point mutations and their effects on binding affinity to a certain target.

Analysis of the interaction of binders with antigens

Before the final clones were titrated against varying concentrations of antigen, preliminary experiments were performed using a fixed amount of antigen (CON3, CON4, CON5, full-length peroxidasin) in order to get a rough estimate of the binding affinities of the different clones. These pre-experiments were done with nine clones termed P1.5.1 to P1.5.9. From those 9 clones 4 were chosen for further affinity maturation (P1.5.1, P1.5.6, P1.5.8 and P1.5.9). After the final selection round 17 affinity-matured, clones were derived from those 4 clones. With all 21 clones a pre-screening was performed using fixed amounts of antigen in order to find the best binders with highest affinity and highest expression levels.

After 3 rounds of bead selection, 1 round of mutagenesis by error-prone PCR and 4 selection rounds using FACS, 9 single clones were selected and further characterized (P1.5.1 - P1.5.9). The selected yeast cell clones were grown in SG-CAA overnight at 20°C in order to express the binder scaffold on their surface. An appropriate amount of cells (OD_{600} of 0.5) was added to 300 µL of PBSA and centrifuged (2000 g, 5 min, 4°C). The PBSA has the effect that the cells pellet more efficiently in the micro-centrifuge tube. Cells were washed with 1 mL PBSA, centrifuged and resuspended to an OD_{600} of 0.5. In a 96-well plate 20 µL of cells were incubated with 20 µL buffer containing 2x concentrated primary reagents per well. Cells were incubated for 1 hour while shaking at 4°C in the dark. After three washing steps with 180 µL PBSA per well (centrifugation and discarding of buffer after each washing step) the cells were resuspended in 40 µL buffer containing the appropriate amount of fluorescently labelled secondary reagents and incubated for 20 minutes at 4°C while shaking. After a final washing step with PBSA the cells were centrifuged. Immediately before flow cytometric analysis the cells were resuspended in 100 - 150 µL of buffer and analyzed on FACS Canto™ II (BD Biosciences).

The first titration curves were obtained using clones P1.5.1, P1.5.2, P1.5.3, P1.5.4, P1.5.5, P1.5.6, P1.5.7, P1.5.8 and P1.5.9 with biotinylated CON4 as antigen and mouse-anti-c-myc antibody (0.5 mg/mL, 1:100) as primary detection reagent. To detect binding affinity to antigen and expression level of binders on the yeast cell surface, cells were labelled with Streptavidin-AF647 (2.0 mg/mL, 1:500) and goat-anti-mouse-AF488 (2.0 mg/mL, 1:500),

respectively, in a second labelling step. Biotinylated CON4 was used in the following concentration range: 1000, 316, 100, 32, 10, 3.2, 1.0 and 0 nM, respectively (0 nM served as blank which was subtracted from all measurements).

In another experiment, we used non-biotinylated CON4, CON5, full-length peroxidasin and cetuximab, with a fixed concentration of 200 nM for all antigens except for cetuximab (1000 nM) in order to estimate binding. The clones used in this experiment were the same as mentioned above (P1.5.1, P1.5.2, P1.5.3, P1.5.4, P1.5.5, P1.5.6, P1.5.7, P1.5.8 and P1.5.9). Cells were prepared as described above. 20 µL cells were mixed with 20 µL of the corresponding antigen (CON4, CON5, PXDN, cetuximab). After three washing steps with 180 µL PBSA per well (centrifugation and discarding of buffer after each washing step) 40 µL of fluorescently labelled secondary reagents were added. For PXDN, CON4 and CON5 Penta-His-AF488 (0.5 mg/mL, 1:200) was used to detect binding to the antigen and anti-HA-AF647 (0.5 mg/mL, 1:300) was used to detect the expression levels of the binder. For cetuximab anti-human CH2 antibody was applied as secondary reagent.

In a third experiment, we used CON3, CON4 and again full-length peroxidasin. All antigens were used at a concentration of 200 nM. Cells and fluorescently labelled detection reagents were prepared as described above.

After finishing the entire selection process with 21 obtained single clones another titration with two different concentrations (5 and 100 nM) of non-biotinylated CON4 and 200 nM of full-length peroxidasin was executed in order to select for those binders which bind with highest affinity for the antigen. From these 21 clones 4 were already obtained earlier. From these initial 4 clones (P1.5.1, P1.5.6, P1.5.8 and P1.5.9) the other 17 were derived through the affinity maturation process.

Titration Curves (Triplicates)

In the end 11 clones were chosen based on their binding signal and their expression level determined in prior experiments. For those 11 binders K_D values were determined by titrating with varying concentrations of non-biotinylated CON4. The following concentration steps were used in this experiment: 600, 200, 66.7, 22.2, 7.4, 2.5, 0.8 and 0 nM. Binding and surface expression was detected by using anti Penta-His-AF488 (0.2 mg/mL, 1:100) and anti-HA-AF647 (0.5 mg/mL, 1:300), respectively.

3.3.12. Competition Experiment

In order to investigate if all obtained binders recognize the same or bind to different epitopes a competition experiment was performed. In this experimental setup, yeast cells expressing binders on their surface were mixed with antigen (CON4) pre-incubated with solubly expressed binders. If the binders on the surface and the soluble binder (pre-incubated with antigen) recognize the same epitope no signal will be obtained using FACS. If the soluble binder recognizes a different epitope on the antigen compared to that of the surfaced-expressed mutant, a signal will be obtained. An overview of this procedure is given in Figures 10 and 11.

Yeast cells were induced in order to express binders on their surface. Induced cells were treated like already described in section 3.3.10. After pipetting the cells into the 96-well plate a mixture of antigen (CON4) pre-incubated (10 min, 4°C) with soluble POX binder (POX = peroxidasin) was added. To 25 µL of cells 25 µL of the protein mixture containing binder (1 µM) and CON4 (10 nM) was added and incubated for 30 - 60 min at 4°C shaking. In addition, binding in the absence of soluble binders (i.e. 10 nM CON4 only) was also measured as a reference. After 3 washing steps with PBSA the cells were resuspended in cold buffer containing secondary reagents: Penta-His-AF488 at a 1:100 dilution (stock concentration: 0.2 mg/mL) and Anti-HA-AF647 at a 1:300 dilution (stock concentration: 0.5 mg/mL).

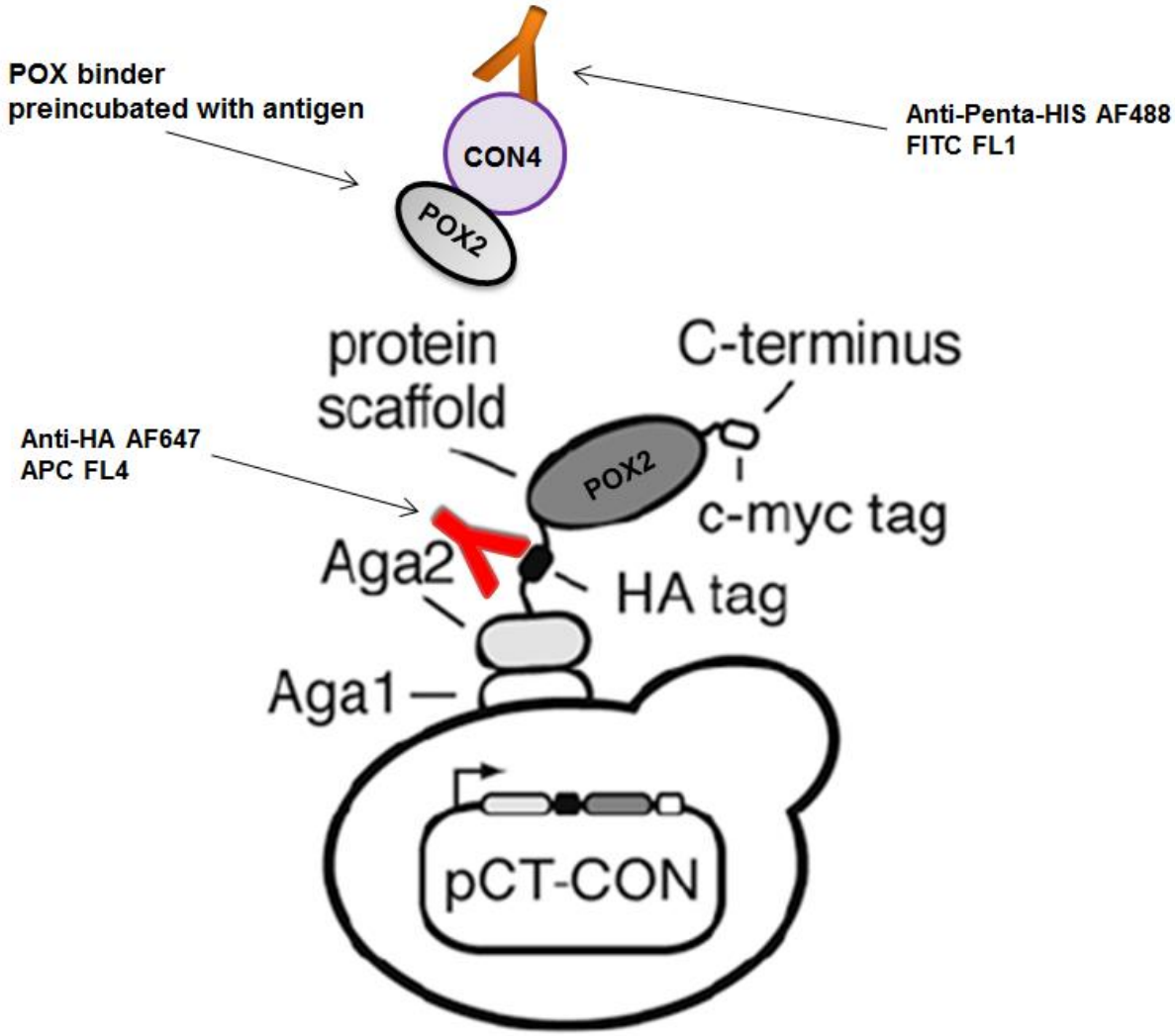


Figure 10. Competitive binding. In this case, the soluble peroxidasin binder POX2 was incubated with antigen CON4. If this mixture is applied to yeast cells expressing the same binder on their surface (or another binder recognizing an overlapping epitope) no signal (or reduced signal) will be obtained due to the fact that the epitope is already blocked. Figure was adapted from: Angelini, Alessandro, et al. "Protein Engineering and Selection Using Yeast Surface Display." Yeast Surface Display: Methods, Protocols, and Applications (2015): 3-36.

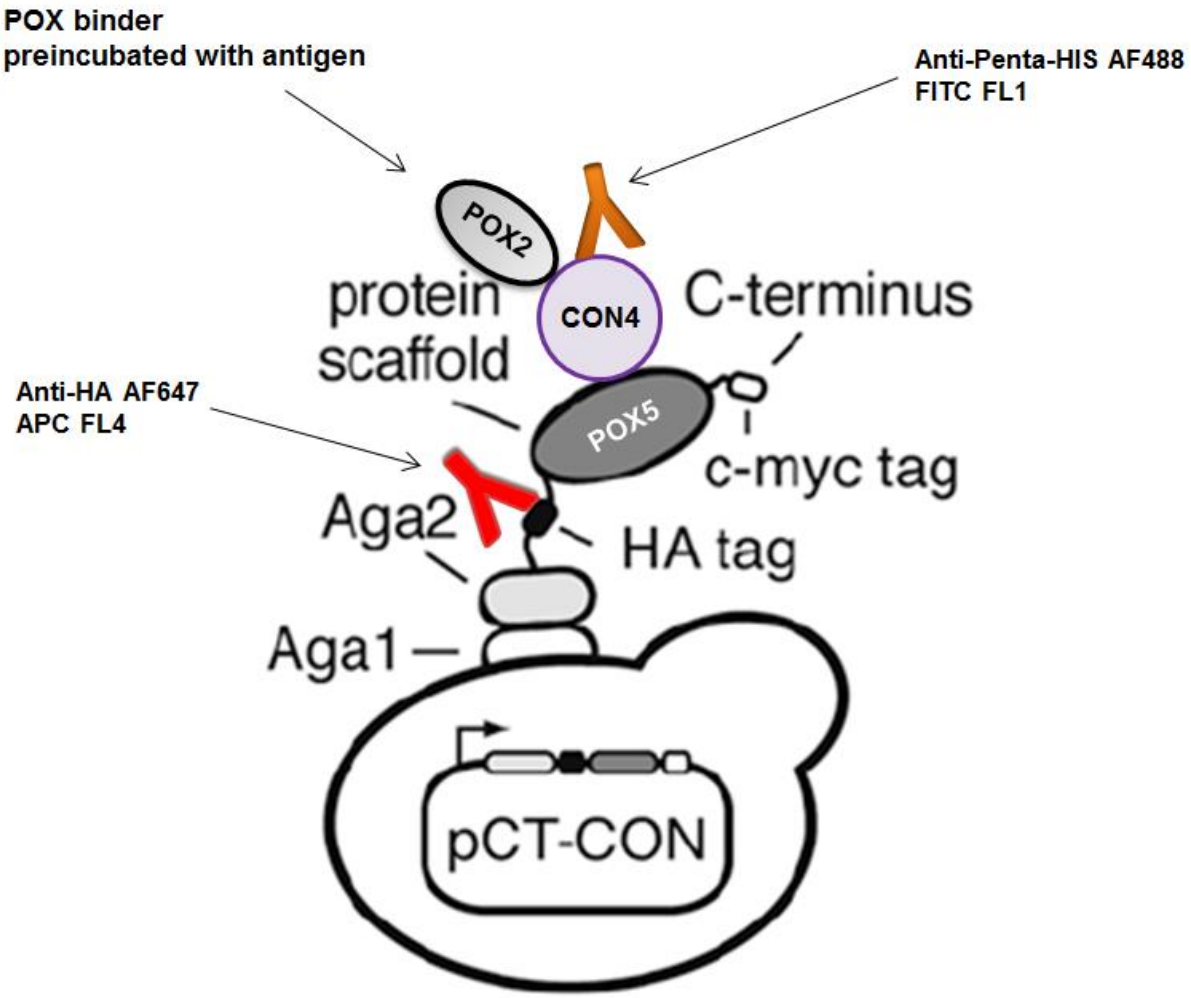


Figure 11. Non-competitive binding. In this case, the soluble peroxidasin binder POX2 was incubated with antigen CON4. If this mixture is applied against yeast cells expressing a non-competitive binder (e.g. POX5) on their surface a signal will be obtained because the epitope is not blocked. Both binders can recognize their antigen simultaneously on different epitopes. Figure was adapted from: Angelini, Alessandro, et al. "Protein Engineering and Selection Using Yeast Surface Display." Yeast Surface Display: Methods, Protocols, and Applications (2015): 3-36.

3.4. SUMO Protease

3.4.1. Expression of SUMO Protease

Inserts cloned into the pE-SUMO expression vector have the advantage (besides the high yield when expressed in *E.coli*) that proteins contain a SUMO-tag and an N-terminal Hexa-His-tag. SUMO protease specifically recognizes the Smt3 (SUMOpro fusion tag) structure and cleaves just before the first amino acid of the protein of interest (without cleaving within the protein of interest).

Already transformed *E.coli* cells (Rosetta DE3) carrying the plasmid for the SUMO protease were incubated in 5 mL LB medium with ampicillin and chloramphenicol as selection marker overnight at 37°C. On the next day 5 mL of the previous culture was inoculated in 1 L of LB (containing the same antibiotics as mentioned before) and were grown till an OD_{600} of 0.6 - 1.0 was reached. After induction with 1 mM IPTG the culture was incubated overnight at 20°C on an orbital shaker.

3.4.2. Purification of SUMO Protease

The cell broth was spun down for 20 minutes (4250 g, 4°C). The pellet was resuspended in 50 mL of resuspension buffer and sonicated for 60 seconds 3 times (duty cycle 50 %, amplitude set to 5, always on ice). The cells were centrifuged at 20.000 g for 15 minutes at 4°C. The supernatant was recovered and filter sterilized (0.45 µm). A 1 M imidazole buffer was added to reach a final imidazole concentration of 1 mM (42 µL added to 42 mL supernatant).

Mini-chromatographic columns (diameter of 1.5 cm) were filled with 2 mL TALON® resin and washed with 50 mL ddH_2O followed by 50 mL equilibration buffer. The supernatant of the expressed protein was added in 3 steps á 14 mL. The whole protein solution was flown 2 times through the column to ensure more complete binding of the tagged protein to the resin. The resin was washed twice with 20 mL equilibration buffer containing 5 mM and 15 mM imidazole, respectively. Finally, proteins were eluted with 2x 5 mL of equilibration buffer containing 250 mM imidazole. All fractions were collected (flow-through, 5 mM, 15 mM, eluate and waste). Afterwards the resin was washed with 2x 20 mL 250 mM imidazole equilibration buffer, 3x 25 mL regeneration buffer, 100 mL ddH2O and 50 mL of 20 % ethanol. The resin was resuspended in 20 % ethanol and stored at 4°C for future use.

3.4.3. Validation, Quantitation and Storage

To separate molecules according to their molecular weight proteins are subjected to an electrical field which allows them to migrate through a polymerized mesh-like gel. Sodium-dodecyl-sulfate (SDS) is an anionic detergent which is used to denature proteins and which confers negative charge over the whole protein in proportion to its molar mass. This allows separation of proteins by their size. Here SDS-PAGE was used to elucidate the purity of proteins.

To validate the purity and size of expressed SUMO protease a SDS-PAGE was performed using a 4-12 % Tris acrylamide 15-well gel (Novagen). 10 µL of each fraction in different dilutions (with ddH2O and 4xSDS reducing buffer) was loaded on the gel after heating all samples for 5 min at 95°C. Mark12™ (Invitrogen) served as marker. The samples were run for 30 min at 200 V.

The gels were rinsed with distilled water and stained with Coomassie Blue R-250 at room temperature for 30 minutes on a shaker. After staining the gels were washed with distilled water and destained until bands could be seen with a clear background (20 - 40 min).

Table 6. Components used for coomassie staining.

Staining solution	**Destaining solution**
50 % v/v distilled water	50 % v/v distilled water
0.1 % w/v coomassie Blue R-250	10 % v/v acetic acid
10 % v/v acetic acid	40 % v/v methanol
40 % v/v methanol	

Next, the proteins were concentrated and buffer-exchanged (Equilibration buffer) 2 times to a final volume of 5 mL using Amicon 3K unltrafiltration devices. The protein concentration was measured at 280 nm. Proteins were dialysed against SUMO-storage buffer at 4°C in 3 steps with 5.0 %, 7.5 % and 10.0 % glycerol, respectively. All dialyzation steps were carried out at 4°C for at least 4 hours per glycerol concentration. The protein concentration was measured and finally SUMO protease was adjusted with 100 % glycerol to a final concentration of 50 %, aliquoted á 400 µL and stored at -80°C.

3.5. Expression of Binders

3.5.1. Sub-cloning into SUMO Expression Vector

Cells tag proteins ready for degradation with ubiquitin, an 8.5 kDa protein and target them to the 26S proteasome. In contrast, the small ubiquitin-like modifier (SUMO) regulates cell processes like nuclear transport, signal transduction and protein stabilization. SUMO is conserved (yeast to humans) and occurs in most eukaryotic systems.

The fusion of yeast SUMO (Smt3) with a protein of interest often enhances expression, promotes correct folding (chaperoning effect) and solubility. Therefore, the attachment of SUMO to a protein of interest is highly desirable for the process of recombinant protein expression. A further advantage of the SUMO expression system is the relatively easy removal of the tag after expression. SUMO protease recognizes the Smt3 (SUMOpro fusion tag) structure at the N-terminus of the fused protein of interest and cleaves at this site without cleaving within the protein of interest.

The pE-SUMO vector allows regulation compatible with the pET expression system. The plasmid contains resistance to kanamycin. The vector is provided as circular plasmid. For cloning purposes, the vector has to be digested with *Bsa*I endonuclease, a class IIS restriction enzyme which recognizes a non-palindromic sequence.

For the purpose of sub-cloning of 11 selected clones from the pCT-CON2 vector to the pE-SUMO expression vector the SUMOpro® Gene Fusion Technology kit was used and primers were designed according to the protocol. All PCR amplified inserts were digested with restriction endonucleases *Xba*I and *Bsa*I, thereby generating fragments having complementary sticky ends to the expression vector linearized with *Bsa*I.

The example of a forward primer design containing a *Bsa*I Class IIS restriction site is given below:

*Bsa*I 5’-NN **GGTCTCN<u>AGGT</u>** XXX NNN NNN NNN NNN NNN-3’

GGTCTC is the recognition sequence of *Bsa*I, N is any nucleotide. The overhang generated by *Bsa*I will be AGGT.

An example of a reverse primer, which should contain one of the restriction sites from the multiple cloning site (MCS) provided by the pE-SUMO vector, is given below:

*Xba*I 5’-NN **TCTAGA** TTA XXX NNN NNN NNN NNN-3’

The restriction enzyme *Xba*I recognizes the sequence TCTAGA, TTA is the reverse complement of the stop codon TAA and XXX is the reverse complement of the final codon of the gene of interest.

3.5.2. Primer Design

The following 9 primers were designed in order to sub-clone the gene of interest from the pCT-CON2 vector to the pE-SUMO expression vector for efficient recombinant expression in *E.coli* (Rosetta DE3):

SUSso_fwd_rcSso
ATTCTGGTCTCTAGGTGCAACCGTGAAATTCACATACCAAGGCGAAGAAAAAC

SUSso_back_rcSso
AAGTATCTAGATTATTGCTTTTCCAGCATCTGCAGCAGTTCTTTC

SUSso_fwd_T2I
ATTCTGGTCTCTAGGTGCAATCGTGAAATTCACATACCAAGGAGAAGAAAAAC

SUSso_fwd_T2V
ATTCTGGTCTCTAGGTGCAGTCGTGAAATTCACATACCAAGGAGAAGAAAAAC

SUSso_fwd_K12R
ATTCTGGTCTCTAGGTGCAACCGTGAAATTCACATACCAAGGCGAAGAAAGAC

SUSso_back_Q56R
AAGTATCTAGATTATTGCTTTTCCAGCATCCGCAGCAGTTCTTTC

SUSso_back_Q61R
AAGTATCTAGATTATCGCTTTTCCAGCATCTGCAGCAGTTCTTTC

SUSso_fwd_T2A_Q8L
ATTCTGGTCTCTAGGTGCAGCCGTGAAATTCACATACCTAGGCGAAGAAAAAC

SUSso_fwd_F5Y_Q8Y
ATTCTGGTCTCTAGGTGCAACCGTGAAATACACATACTACGGCGAAGAAAAAC

Regions annealing to the rcSso7d binders in pCT-CON2 vector are highlighted in yellow. The annealing region of primers was checked for silent mutations and the sequence of the Sso7d gene was checked for potential restriction sites for *Xba*I and *Bsa*I

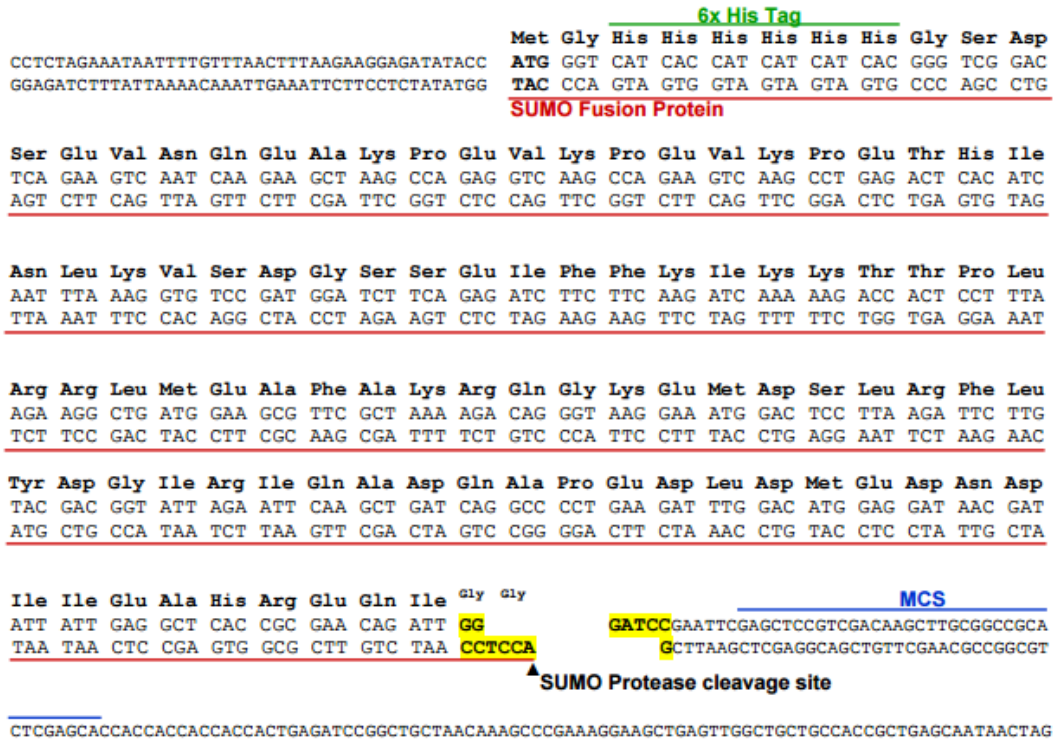


Figure 12. Polylinker map of pE-SUMO vector (5598 bp) used for the expression under the control of a T7 promoter and a kanamycin resistance marker. The SUMO protease cleavage site is highlighted in yellow, MCS is marked in blue, the actual SUMO fusion protein is underlined in red and the 6x His tag in green. Figure was adapted from:

https://www.lifesensors.com/wp-content/uploads/2017/03/1100A_1106_Polylinker_Map_Amp.pdf

3.5.3. Amplification of Template DNA

DNA inserts of potential peroxidasin binders were amplified by PCR using Phusion polymerase, which has an error rate more than 50-fold lower than that of *Taq* DNA polymerase, a tenfold shorter extension time compared to *Pfu* DNA polymerase and high yield. The conditions of the PCR reaction are shown in Tables 7 and 8. An overview of the used primer pairs for each insert is given in Table 9 below.

Table 7. PCR mixture used for amplification of inserts using a high fidelity polymerase.

Component	Volume [µL]
ddH2O	32.5
5x Phusion HF Buffer	10
dNTPs (10 mM each)	1.0
Forward primer (10 µM)	2.5
Reverse primer (10 µM)	2.5
Phusion polymerase (2 U/µL)	0.5
Template DNA (1:100 dilution from miniprep)	1.0
Total Volume	50

Table 8. Cycling condition of the PCR reaction.

Segment	Cycles	Temperature [°C]	Time [sec]
1	1	98	30
2	30	98	10
		72	15
3	1	72	300

Table 9. Primers used for the corresponding insert from the respective clone.

Clone Name	Forward Primer	Reverse Primer
P1.5.1	SUSso_fwd_T2I	SUSso_back_rcSso
P2.3.1.3	SUSso_fwd_T2V	SUSso_back_rcSso
P1.5.6	SUSso_fwd_K12R	SUSso_back_rcSso
P3.3.6.2	SUSso_fwd_K12R	SUSso_back_Q61R
P3.3.6.3	SUSso_fwd_K12R	SUSso_back_Q61R
P1.5.8	SUSso_fwd_rcSso	SUSso_back_rcSso
P3.3.8.2	SUSso_fwd_rcSso	SUSso_back_rcSso
P3.3.8.3	SUSso_fwd_rcSso	SUSso_back_Q56R
P1.5.9	SUSso_fwd_rcSso	SUSso_back_rcSso
P3.3.9.1	SUSso_fwd_T2A_Q8L	SUSso_back_rcSso
P3.3.9.3	SUSso_fwd_F5Y_Q8Y	SUSso_back_rcSso

The amplified inserts were checked by a 2 % agarose gel. The fragments roughly corresponded to their expected size of 213 bp (183 bp (Sso7d) + 30 bp (primers)).

3.5.4. Digestion of Amplification Products

The amplified and purified inserts were digested using *Bsa*I-HF and *Xba*I in combination with CutSmart buffer. The mixture was incubated 2 – 3 hours at 37°C. After another purification step by using the illustra GFX PCR DNA and Gel Band Purification Kit the concentration of the digested products was measured and the fragments were analysed by gel electrophoresis to check the fragments.

Table 10. Reaction mixture of double digest with BsaI-HF and XbaI.

Component	Volume [µL]
Purified PCR-product (insert)	38
CutSmart buffer 10x	4.5
BsaI-HF	1
XbaI	1

3.5.5. Preparation of pE-SUMO Expression Vector

E.coli cells expressing the empty pE-SUMO vector were incubated overnight at 37°C in LB medium containing kanamycin as selection marker. Plasmids were isolated according to a standard miniprep protocol using the QIAprep Spin Miniprep Kit. The vector was digested with *Bsa*I HF in CutSmart buffer for 3 hours at 37°C. After digestion the vector was purified in order to ensure that no more restriction enzyme and corresponding restriction buffer is left. To prevent religation of the empty pE-SUMO vector a CIP-digestion was executed using calf intestinal alkaline phosphatase (CIP) (New England Biolabs®). CIP catalyzes the dephosphorylation of 5’ and 3’ ends which prevents the religation of linearized DNA. The vector was incubated in CutSmart buffer with CIP for 1.5 hours at 37°C. After another purification using the illustra GFX PCR DNA and Gel Band Purification Kit the vector was ready for ligation with the inserts encoding the 11 selected clones.

3.5.6. Ligation of pE-SUMO Vector with Selected Mutants

In order to ligate the inserts of the 11 selected mutants into the expression vector it is necessary to adjust the ratio of digested insert to vector backbone. The molar ratio of insert to vector was set to 3:1. The calculated amounts of vector and insert were mixed together with T4 DNA ligase (catalyzes the formation of phosphodiester bonds between 5’ phosphate and 3’ hydroxyl groups) and T4 ligase buffer and incubated for 2 hours at room temperature.

3.5.7. Transformation into *E.coli* Rosetta DH3

Heat shock competent *E. coli* Rosetta DH3 cells were transformed with 2.5 µL of the ligated and purified pE-SUMO-rcSso7d vector according to the protocol described in section 3.3.9. 55 µL as well as ~ 100 µL of concentrated cells were plated on LB plates containing kanamycin. To concentrate the cells they were centrifuged for 1 minute at full speed (16000 rpm). All plates were incubated at 37°C overnight.

3.5.8. Miniprep and Sequencing of Binder DNA in pE-SUMO

All inserts introduced into the pE-SUMO expression vector and a 36 bp sequence before the SUMO cleavage site were checked for mutations. All sequencing steps were carried out by Microsynth AG using Sanger sequencing technology. For every mutant single colonies were picked from KAN-LB plates containing 50 µg/mL kanamycin (KAN) and inoculated in 5 mL LB medium containing 50 µg/mL kanamycin and 25 µg/mL chloramphenicol (CAM) overnight at 37°C, 180 rpm. The plasmids were isolated using the Illustra™ plasmidPrep Mini Spin Kit from GE Healthcare according to the manufacturer's protocol except for the elution step (HQ-H_2O was used instead of elution buffer). The following primer was used for sequencing (provided by Microsynth AG):

SUMO-5 5'- CCTTAAGATTCTTGTACGACG-3'

From all sequence-checked clones cryo-stocks of overnight cultures were prepared and stored in LB-medium with 30 % glycerol, kanamycin and chloramphenicol at -80°C.

3.5.9. Binder Expression

Two variants of POX-binders were prepared. One version of binders only comprised the protein scaffold rcSso7d. Another version of binders additionally contained two tags: a hexa-histidine tag as well as a SUMO-tag. SUMO protease recognizes a certain structure (Smt3, SUMOpro fusion tag) and cleaves the junction just before the N-terminus of the protein of interest. Therefore, the His-tag and the SUMO-tag can easily be cleaved off in one single step as described in section 3.5.1.
First, 15 µL of a previously used culture (see section 3.3.8) carrying the pE-SUMO plasmid with insert of interest were incubated in a total volume of 5 mL LB medium containing antibiotics (kanamycin and chloramphenicol) overnight at 37°C shaking with 180 rpm.
On the next day, 4 mL of the overnight culture were inoculated in 400 mL terrific broth (TB) with kanamycin and chloramphenicol. After reaching an OD_{600} of 2 expression was induced with 1 mM IPTG (Isopropyl-β-D-1-thiogalactopyranoside) and incubated overnight at 20°C while shaking.

3.5.10. Binder Purification

Cultures were centrifuged (5.000 g, 20 min, 4°C) and the supernatant was discarded. The remaining pellet was resuspended in 30 mL sonication buffer. The cells were disrupted using a sonicator (2x 90 sec, duty cycle 50 %, amplitude set to 5). The lysed cells were centrifuged (20.000 g, 30 min, 4°C) and supernatants filtered through a 0.45 µm filter device.

The filtered supernatants were applied to glass columns (1.5 cm diameter) containing 1.5 mL of TALON® resin. These resins are suitable for batch and low-pressure chromatographic applications. The resin consists of a sepharose with large pore size and a coupled tetradentate chelator charged with Co^{2+} ions which lead to high affinity and specificity for His-tagged proteins. For each binder a separate column was used.

First, the resin was washed with 20 mL equilibration buffer. The samples (containing 10 mM imidazole in order to reduce unspecific binding) were loaded onto the resin twice to ensure that most His-tagged binders were bound to the column. Now the columns were washed twice with equilibration buffer containing 5 mM and 15 mM imidazole, respectively. All samples were eluted twice with 5 mL of equilibration buffer (250 mM imidazole). The used resins were washed 3 times with 20 mL regeneration buffer, 2 times with 20 mL ddH_2O and with 20 mL 20 % ethanol and finally stored in 20 % ethanol at 4°C (long term storage) or at RT for further applications. Protein concentration of the eluate was measured by analyzing the absorption at 280 nm using the NanoDrop.

In order to concentrate and exchange the buffer (from elution buffer to 1xPBS buffer) the eluates of all samples were applied to AMICON® 3K ultrafiltration devices and centrifuged (10 - 20 min, 4°C, 3500 g).

Sample volume of approximately 10 mL was loaded on each ultrafiltration device and centrifuged until 1.0 - 1.5 mL were left. Then the sample was filled up with freshly made PBS buffer and centrifuged again. In total 3 times ~ 10 mL PBS was loaded onto 1.0 - 1.5 mL of protein solution. After ultrafiltration protein concentration was measured.

A fraction of the His-tagged SUMO-fusion binders were frozen in liquid nitrogen in small aliquots and stored at -80°C. To the remaining tagged proteins SUMO-protease was added in a mass ratio of 1:100 (e.g. 1 mg protease per 100 mg SUMO-fusion protein) for an overnight digest at room temperature to cut off the His6-SUMO tag.

On the next day the SUMO-protease-digested samples were applied to TALON resin affinity chromatography as described before. In this purification step the cut binder is found in the flow-through whereas the His-SUMO-tag binds to the column. Prior to loading the resin was split into two parts: ~ 2/3 were left in the column and 1/3 was stored for later usage in order to increase the purification efficiency. Samples were loaded twice on the 2/3 of the resin. Resin was washed two times with 5 mL equilibration buffer not containing any imidazole. Now the 2/3 of the resin were taken out and the remaining 1/3 of the resin loaded on the column. The sample was loaded on the column and flow-through was collected. After washing with 2 times 5 mL equilibration buffer all fractions (flow-through and washing steps from 2/3 of the resin and 1/3 resin) were pooled to yield the purified rcSso7d-based binders. Approximately 12 - 14 mL were collected per binder.

Protein concentration of all samples was determined before the proteins were concentrated and buffer-exchanged using the AMICON-3K ultrafiltration device. Each resin/column was regenerated with 20 mL 250 mM imidazole containing equilibration buffer, 20 mL regeneration buffer, 20 mL ddH2O (3 times each step) and finally washed with 20 mL 20 % ethanol once.

3.5.11. Analysis of Purity by SDS-PAGE

A Novex® NuPAGE® SDS-PAGE Gel System with a NuPAGE® 12 % Bis/Tris gel together with MES-SDS Running buffer was used. From all expressed binders 3 µL of protein were mixed with 12 µL ddH_2O and 5 µL LDS-sample buffer and heated for 5 min on 95°C. Mark 12™ served as a standard. 800 - 900 mL of buffer were filled into the inner and outer chamber of the gel electrophoresis system. 10 µL of each sample was applied together with Mark12™ onto the gel. The gels were run for 30 min at constant voltage of 180 V.

3.6. Biophysical Characterization of Peroxidasin Binders

To further characterize the 11 expressed peroxidasin binders size exclusion chromatography (SEC) and differential scanning calorimetry (DSC) were performed.

3.6.1. Size Exclusion Chromatography

Size exclusion chromatography uses a non-ionic, hydrophilic stationary phase together with an aqueous mobile phase to separate molecules according to their size and shape. Smaller molecules elute later from the column due their longer retention in the porous column material in comparison to larger molecules, which are less retained in the pores.

SEC was used to determine the interaction of binders with the column matrix (indicated by prolonged elution times), which can be used as a metric for non-specific stickiness. Further it was determined if binders elute as monomers or tend to form dimers or higher molecular weight aggregates.

For SEC, 25 µg of protein were injected to a Superdex 200 10/300 column (HPLC system Shimadzu LC-20AD XR). Aliquots of all 11 POX-binders were thawed and diluted accordingly. All samples were centrifuged and filtered before injection with Ultrafree-MC centrifugal devices (3 min, 12.000 g, RT). As a running buffer served SEC buffer (PBS with 200 mM NaCl) , the flow rate was 0.75 mL/min.

3.6.2. Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) is a technique used for the measurement of thermodynamic properties of biomolecules. DSC measures the heat-uptake or release of a sample solution (sample cell) in comparison to buffer without sample (reference cell) during a temperature ramp. This means that both cells are heated up simultaneously and the difference in energy input to maintain equal temperature relates to different heat capacities of the solutions. This provides information on thermally-induced processes in the sample[4].

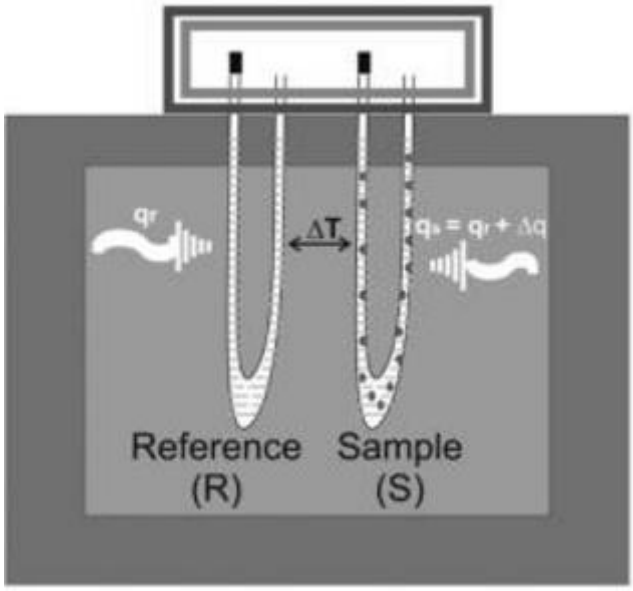


Figure 13. Basic principle of DSC. The figure shows the two cells within a calorimeter: a sample and a reference cell. Both cells are heated up equally. The difference in the amount of energy necessary to maintain the same temperature in both cells gives information on the sample's thermodynamic properties. (Figure adapted from: Gill, Moghadam, Ranjbar, 2010)

The thermal stability of all eleven expressed binders was measured. For that purpose, all POX-binders were diluted to a concentration of 35 µM in PBS in a total volume of 400 µL. Furthermore, a combination of all POX-binders with CON4 was measured. In this set-up 10 µM binders were mixed with 5 µM CON4. All samples were analyzed in PBS. As a blank PBS buffer and a mixture of PBS/SEC buffer was measured. The mixture was used because CON4 was stored in SEC buffer. All scans were performed in a temperature range between 20°C and 120°C. The scan rate was set to 60 °C/h. A buffer baseline was obtained by using a rescan with the same settings or by measuring buffer against buffer to obtain a baseline for the binder measurements. The baseline was subtracted, all concentrations were normalized and the data were fitted with a non-2-state thermal unfolding model.

[4] Gill P, Moghadam TT, Ranjbar B. Differential Scanning Calorimetry Techniques: Applications in Biology and Nanoscience. Journal of Biomolecular Techniques : JBT. 2010;21(4):167-193.

4. Results

4.1. Preparation of Antigens

4.1.1. Size Exclusion Chromatography (SEC)

The antigens used for binder selection (CON4, CON4-B, CON5) and subsequent binder characterization were purified using SEC. This purification step depleted dimers and aggregated proteins.

The chromatograms of CON4, CON4-B and CON5 are shown in Figures 14 – 16 below.

The absorptions of the 3 separations are shown in Table 11. For CON4 fractions 19 – 28, for biotinylated CON4 fractions 19 – 27 and for CON5 fractions 21 – 29 were collected and the concentrations of the samples were measured spectrophotometrically at 280 and 410 nm using the NanoDrop.

Table 11. Absorption at 280 and 410 nm of collected fractions of CON4, CON4-B and CON5 using SEC. White fractions were pooled and used for further experiments.

Fraction	**CON4 (no biotin)**		**Fraction**	**CON4 (with biotin)**		**Fraction**	**CON5 (no biotin)**	
[0.5 mL]	**280 nm**	**410 nm**	[0.5 mL]	**280 nm**	**410 nm**	[0.5 mL]	**280 nm**	**410 nm**
19	0,013	0,004	**19**	0,020	0,009	**21**	0,043	0,007
20	0,015	0,005	**20**	0,020	0,010	**22**	0,005	0,002
21	0,019	0,005	**21**	0,021	0,007	**23**	0,006	0,001
22	0,033	0,011	**22**	0,047	0,018	**24**	-0,001	-0,001
23	0,090	0,028	**23**	0,120	0,041	**25**	0,005	0,003
24	0,123	0,040	**24**	0,144	0,047	**26**	0,010	0,007
25	0,111	0,032	**25**	0,113	0,032	**27**	0,007	0,005
26	0,071	0,020	**26**	0,064	0,020	**28**	0,015	0,009
27	0,046	0,013	**27**	0,047	0,014	**29**	0,003	0,006
28	0,032	0,009						

For CON4 and CON4-B fractions 23 – 26 were pooled whereas for CON5 fractions 27 and 28 (fraction with higher absorption values) and fractions 26 and 29 were pooled (fractions with lower absorption values), respectively. The pooled fractions were measured again using the NanoDrop. The results are shown in Table 12.

Table 12. Absorption of pooled fractions of CON4, CON4-B and CON5.

Sample	Absorption at 280 nm	Absorption at 410 nm
CON4	0.088	0.026
CON4-B	0.104	0.032
CON5-high	0.006	0.003
CON5-low	0.029	0.010

The concentration for CON4 and CON4-B was calculated using Lambert-Beer's law. A is the absorption, c is the concentration and d is the path length of the cuvette in centimetres. The molar extinction coefficient at 280 nm for CON4 is ε = 147500 M^{-1} cm^{-1}.

$$c\ (\text{CON4}) = \frac{A}{(\varepsilon * d)} = \frac{0.088}{147500 * 0.1} = 5.97 * 10^{-6}\text{M} = 5.97\ \mu\text{M}$$

$$c\ (\text{CON4-B}) = \frac{A}{(\varepsilon * d)} = \frac{0.104}{147500 * 0.1} = 7.05 * 10^{-6}\text{M} = 7.05\ \mu\text{M}$$

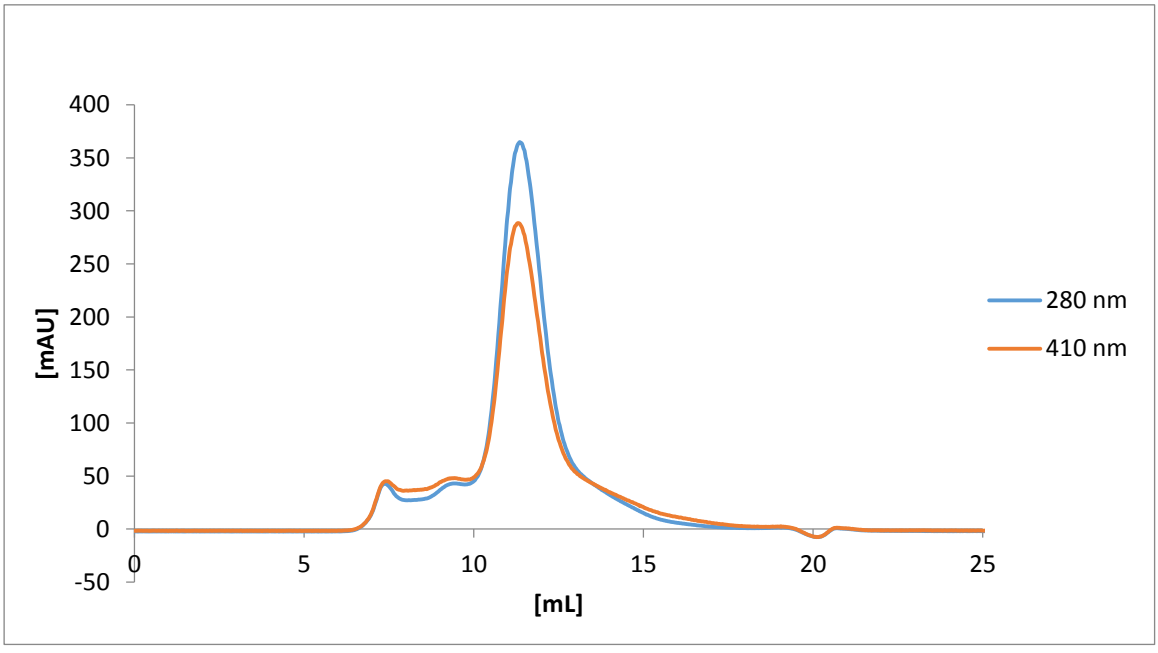


Figure 14. Preparative SEC run of non-biotinylated CON4. 500 µl CON4 were injected onto a Superdex 200 HR 10/300 column at a flow rate of 0.4 ml/min. Absorbance was detected at 280 nm (total protein) and at 410 nm (heme absorbance of CON4). The monomeric CON4 protein eluted at around 11.5 min, whereas oligomers or higher order aggregates eluted earlier. Fractions covering the monomeric peak were pooled and used for further experiments.

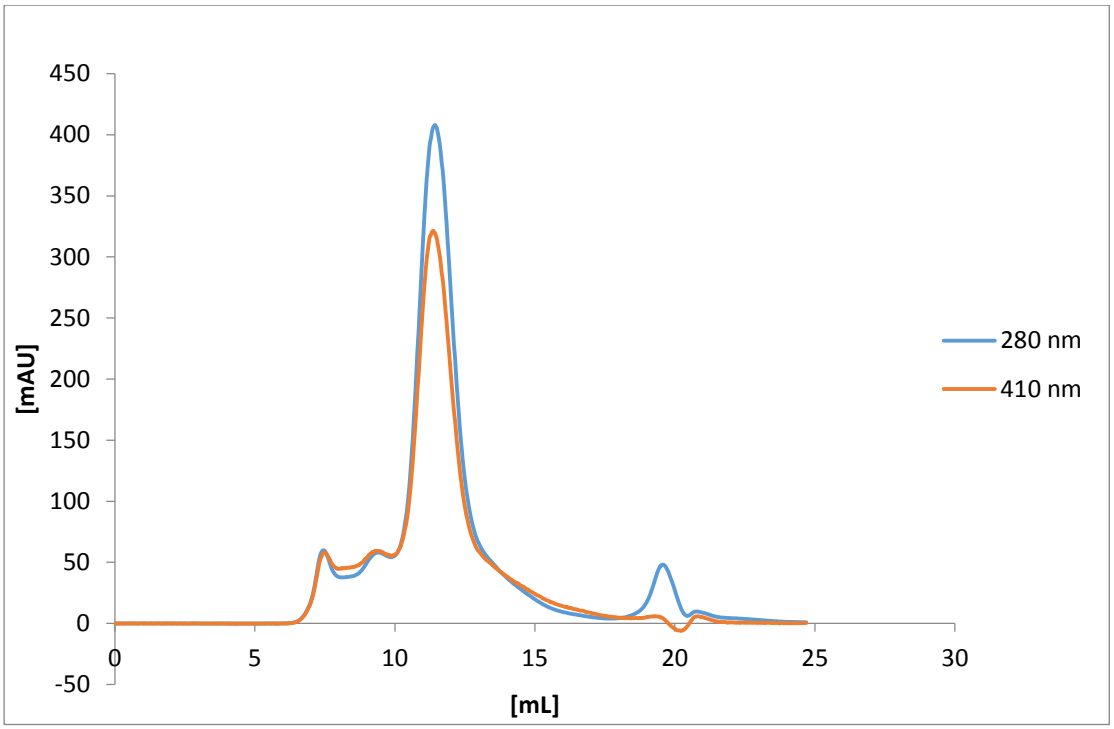


Figure 15. Biotinylated CON4 (CON4-B) in a preparative SEC run on a Superdex 200 HR 10/300 column at a flow rate of 0.4 ml/min. Free biotin elutes at around 19.5 mL (280 nm-line in blue). Total protein absorbance was detected at 280 nm and the heme peak absorbance of CON4-B at 410 nm.

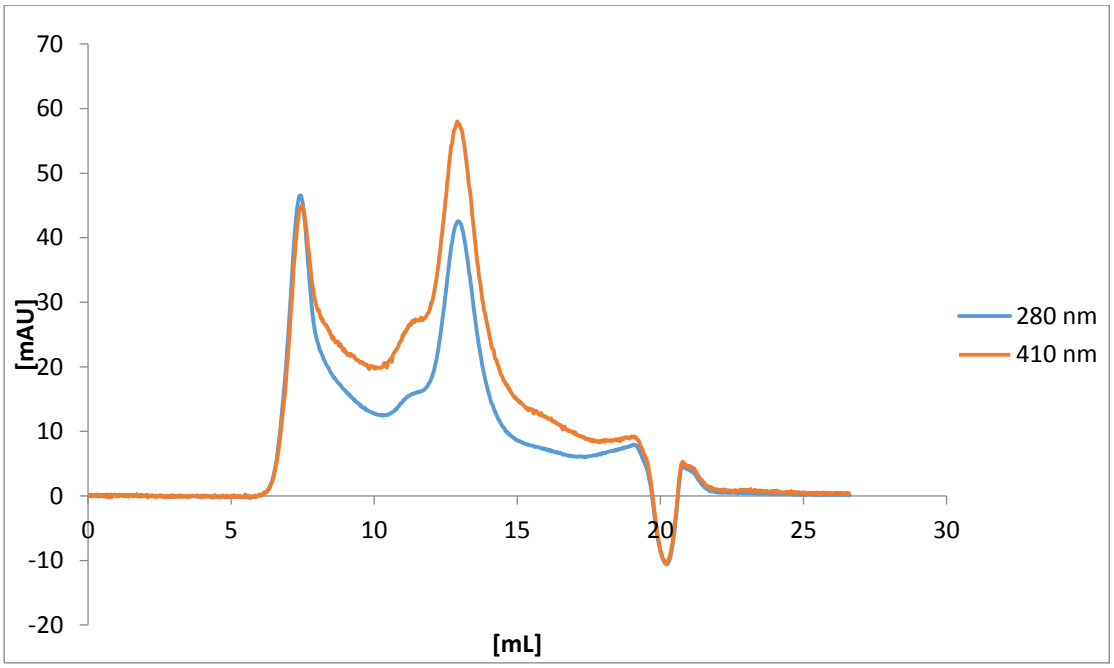


Figure 16. Chromatogram of CON5. The absorbance for two wavelengths is shown (280 nm and 410 nm). The same conditions as mentioned in Figure 15 were applied. Fractions containing the monomeric peak (eluting at around 13 ml) were collected and used for further experiments.

4.1.2. Verification of Biotinylation of CON4

SEC-HPLC-MALS was used to determine if the antigen CON4 was successfully biotinylated.

Three samples were prepared as shown in Table 13 either only CON4-B, or CON4-B mixed with streptavidin-AF647 in a 3-fold molar excess, or streptavidin-AF647 only.

Table 13: Samples prepared for SEC-HPLC-MALS.

	CON4-B	**CON4-B + Streptavidin**	**Streptavidin**
CON4-B	846 µg/mL	846 µg/mL	0 µg/mL
Streptavidin	0 µg/mL	2000 µg/mL	2000 µg/mL

The AF647-labelled streptavidin, which could be traced at 650 nm, enabled the identification of peaks that contained streptavidin. Due to the binding of streptavidin to biotin a peak shift occurs (shoulder around 13.5 min). Aggregates elute from the column after around 11.5 min. Free streptavidin elutes at 19.5 min. A direct comparison of CON4-B and CON4-B bound to streptavidin can be found in Figure 17.

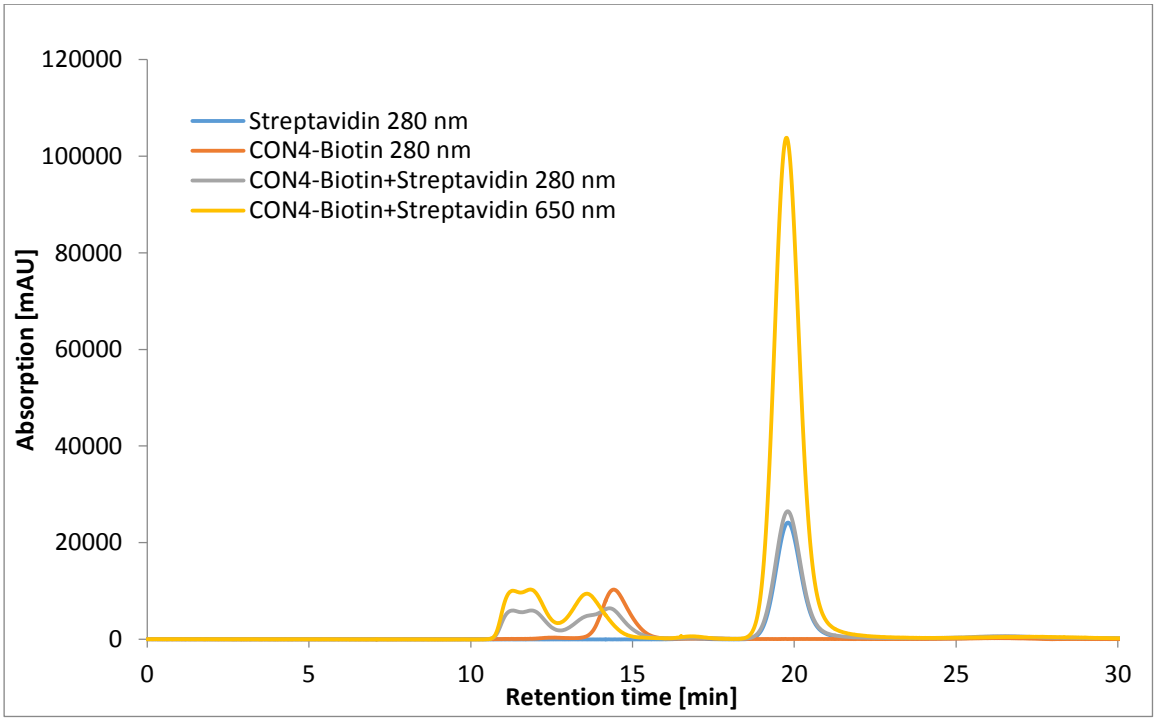


Figure 17. Chromatogram of the SEC analysis of CON4-B, CON4-B with added streptavidin and streptavidin only. Streptavidin was labelled with AF647 which allowed tracing the absorbance at 650 nm in order to identify peaks that contained streptavidin. The peaks appearing after approximately 20 min are streptavidin peaks. The curve in blue shows only streptavidin absorbance detected at 280 nm, the yellow and the grey curve show CON4-B bound to streptavidin. The orange curve shows CON4-B, the biotinylated version of CON4, which appears after ca. 14 min.

4.2. Results of the Selection Process

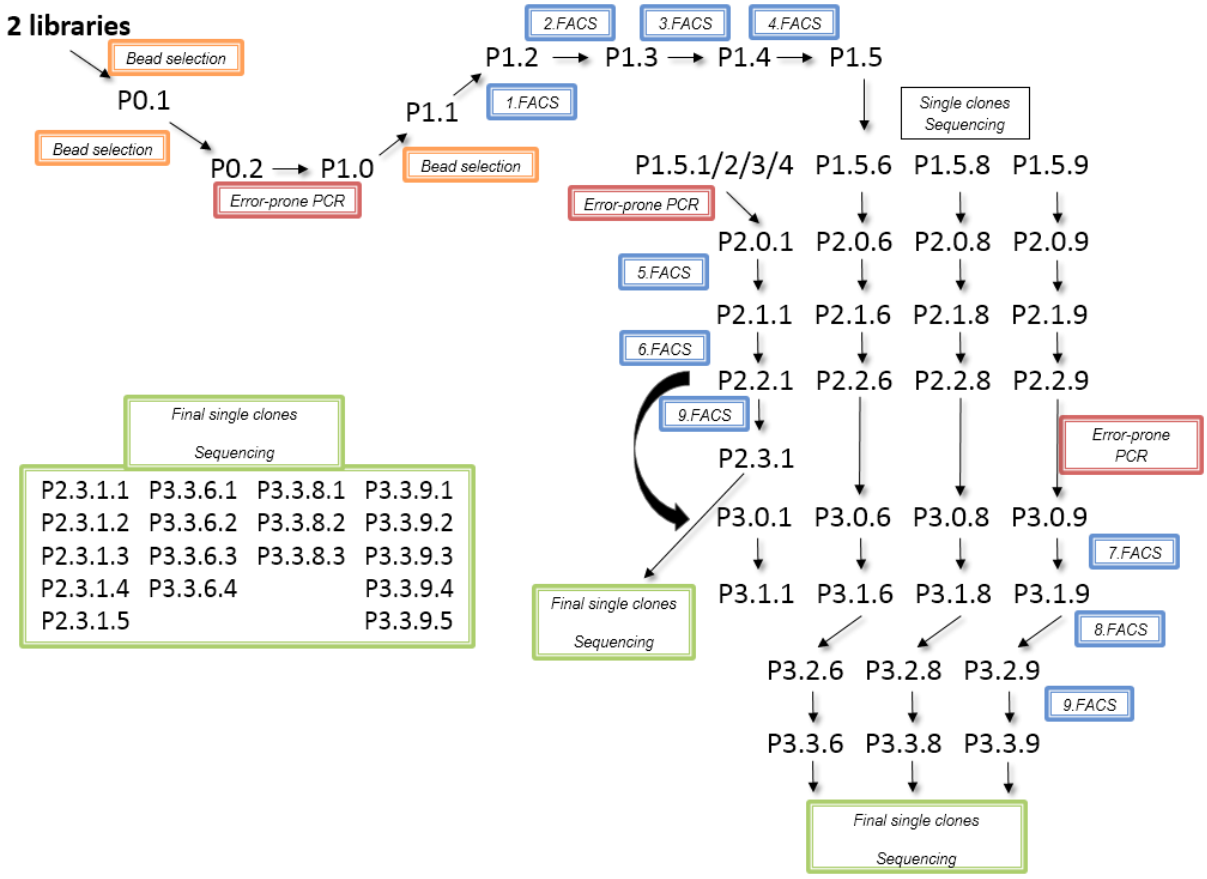


Figure 18. Detailed overview of the selection process. Starting from two libraries (rcSso7d-11 and rcSso7d-18) 3 rounds of magnetic bead selection, one round of mutagenesis by error-prone PCR and 4 FACS selections were executed before the first single clones were isolated. After additional rounds of mutagenesis and FACS selections the final clones were used for determination of binding affinity via yeast surface titrations in triplicates. Furthermore solubly expressed binders were biophysically characterized (differential scanning calorimetry (DSC), size exclusion chromatography (SEC)).

4.2.1. Magnetic Bead Selection

In order to facilitate selection of binders even with low affinity, magnetic bead selection was performed 3 times in total. Starting from the initial libraries rcSso7d-11 and rcSso7d-18, two selection rounds (including two positive and three negative selections using antigen-coated or bare beads, respectively) were executed consecutively. Another magnetic bead selection round (including positive and negative selections) was done directly after the first error-prone PCR.

The library diversity of both libraries (rcSso11 and rcSso18, which were pooled later) was 1.4×10^9 clones each. After the first bead selection (library P0.1) a library diversity of 1.2×10^7 clones was obtained. After the second magnetic bead selection two libraries with diversities of 5.6×10^5 (P0.2) and 2.0×10^5 (P0.2 depleted) clones were gained, respectively. Before the next bead selection was executed, the libraries were mutagenized by using error-prone PCR (yielding libraries P1.0 and P1.0 depleted). With these two mutagenized libraries the third and last bead selection led to a diversity of 8.2×10^5 (P1.1) and 1.0×10^6 (P1.1 depl.) clones.

4.2.2. Mutagenesis of Libraries by Error-prone PCR

Randomization of libraries containing potential peroxidasin binders was achieved by mutagenesis using error-prone PCR.

In total three rounds of error-prone PCR were executed. The first one was executed after 2 rounds of magnetic bead selection. The next one was executed after the first single clones were sequenced (clones P1.5.1 - P1.5.9) and a last error-prone PCR after the 6th FACS sort leading to libraries P3.0.1, P3.0.6, P3.0.8 and P3.0.9 (schematically shown in Figure 19).

4.2.3. Fluorescent Activated Cell Sorting

Fluorescent activated cell sorting was used to select high-affinity binders. Preparation of the yeast cells carrying potential binders was performed as described in section 3.3.10. In the initial rounds of the selection process around 1 - 3 % of the cells were collected to avoid loss of diversity. In later rounds only around 0.1 - 0.5 % of the sorted cells were selected - the selection process in these rounds is more stringent meaning that only the binders with the highest affinity and the cells with the best expression rates are selected.

The first FACS round was performed with libraries P1.1 and P1.1 depl. (after 3 bead selection rounds and one round of mutagenesis by error-prone PCR between the 2nd and 3rd bead selection round). Four plots of such a FACS selection round (library P2.0.6 and P3.0.8) are shown exemplarily in Figure 19. On the X-axis the display level (expression of binders) is shown. The population in the lower left corner of the dot plots indicates yeast cells which do not express the binder, which is characteristic for yeast display (Angelini et al., Methods Mol. Biol. 2015). Binding to the antigen is shown on the Y-axis. The two charts on the left side show negative controls (without antigen). The library P3.0.8 shows an increased slope of the diagonal distribution of the yeast cell population than library P2.0.6, an indication of higher affinity towards the antigen.

A detailed overview of used antigens, detection reagents and number of cells per library can be found in Table 5 in section 3.3.10.

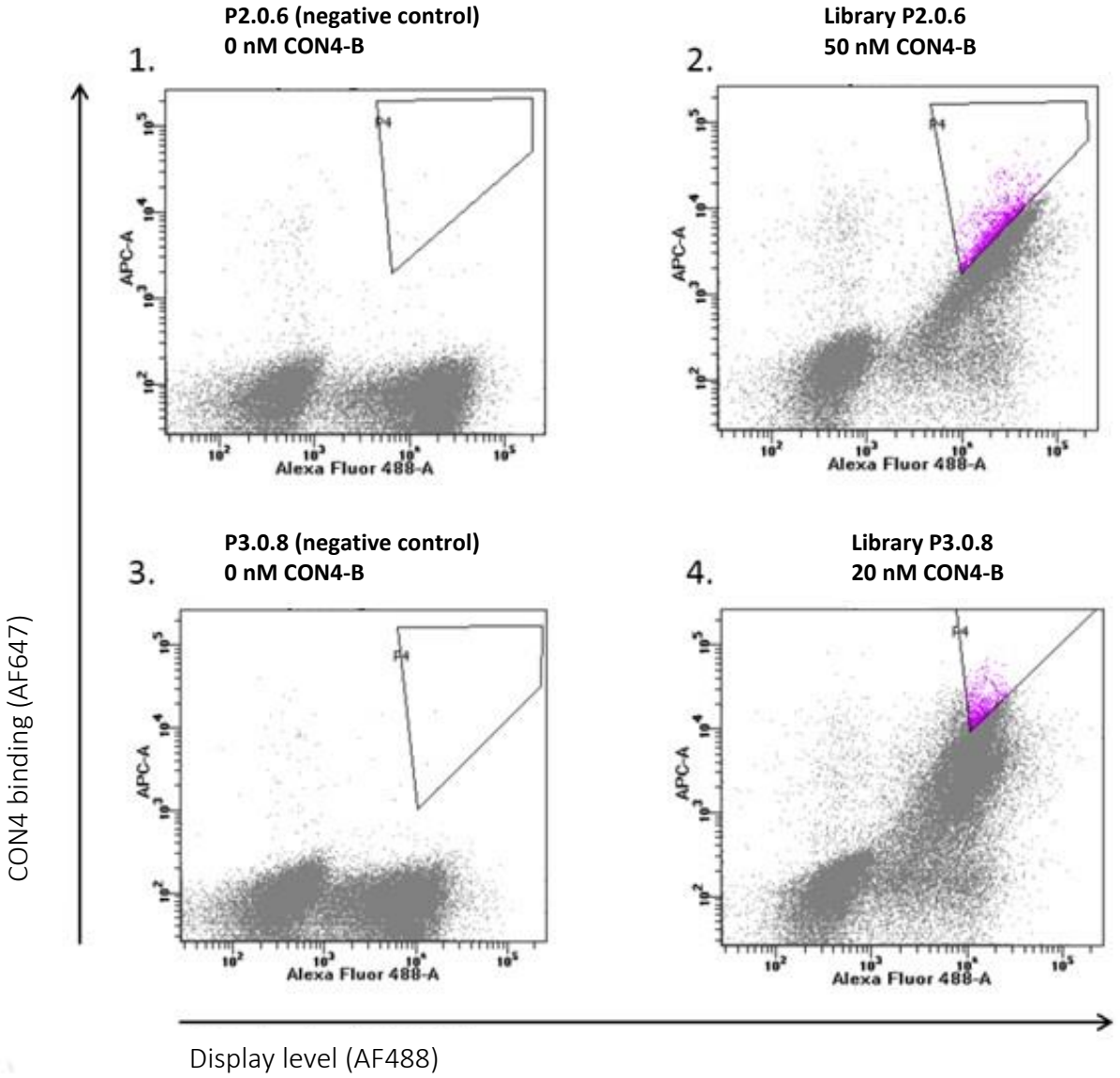


Figure 19. Fluorescent activated cell sorting of two different libraries (P2.0.6 from the 5th FACS selection round and P3.0.8 from the 7th FACS selection round) against biotinylated antigen CON4 with corresponding negative controls. Dot plots 1 and 2 show the yeast cell population of library P2.0.6 whereas charts 3 and 4 show the yeast cell population of library P3.0.8. The two charts on the left (1 and 3) are the negative controls without added antigen. The two left charts (1 and 3) show the negative controls (no antigen added). There are two yeast cell populations in those two negative controls: The one in the bottom left corner is a non-expressing one (x-axis is indicating expression levels). The second population of cells which can be seen in these two charts are cells which express potential binders on their surface but due to the fact that no antigen was added no binding can occur. A direct comparison of these plots is not possible because two different concentrations of CON4-B was used.

4.3. Binder Characterization

4.3.1. Amino Acid Sequences of Binders

After two rounds of magnetic bead selection (in order to retain weak affinity binders), one round of error-prone PCR and another round of a magnetic bead selection, four consecutive rounds of FACS enrichments were carried out for further selection of high-affinity binders. From library P1.5 nine single clones were obtained and analysed. They were grouped into four binder families according to their amino acid sequence similarities within the engineered binding site of rcSso7d (9 amino acids within the triple stranded beta-sheet)(Figure 20). Family I contains clones P1.5.1 - P1.5.5, family II P1.5.6 and P1.5.7, family III P1.5.8 and family IV P1.5.9. The amino acids of the binding site are at positions 21, 23, 25, 28, 30, 32, 40, 42 and 44 separated by two loops at positions 26 - 27 and 33 – 39, respectively. In all families except for family IV (P1.5.9) single point mutations were found outside the binding site, probably introduced during the epPCR step. An additional amino acid was inserted at position 28 (glutamate) in family I. Within the two loops single point mutations occurred in binder P1.5.3 (E35K), P1.5.6 (Q27H), P1.5.8 (G38S) and in P1.5.9 (G38A).

The theoretical net charge of the protein was determined by summation of all charged amino acids (number of positively charged amino acids – number of negatively charged amino acids = theoretical net charge of protein). In addition, the number of aromatic amino acids within the binding site was also determined. The results are summarized in Table 14.

To further improve the affinities of those 4 sequence families, a mixture of P1.5.1/2/3/4, as well as the individual clones P1.5.6, P1.5.8 and P1.5.9 were further mutagenized by error prone PCR. In the end out of the four binder-families 17 single clones were sequenced (result not shown) and 7 were chosen for further biophysical characterization. In total, we ended up with 11 final clones/binders. The amino acid sequences of the binders (parental clone and affinity-matured variants) are shown in Figure 21 grouped again into four families. Table 15 shows information about the number of aromatic amino acids within the binding site as well as the theoretical net charge of the protein.

The nomenclature was changed for reasons of better readability after the selection process was completed and the final clones expressed, i.e. P1.5.1 has been renamed to POX1, P2.3.1.3 = POX2, P1.5.6 = POX3, P3.3.6.2 = POX4, P3.3.6.3 = POX5, P1.5.8 = POX6, P3.3.8.2 = POX7, P3.3.8.2 = POX8, P1.5.9 = POX9, P3.3.9.1 = POX10 and P3.3.9.3 = POX11. In Figure 21 both names of the corresponding binders can be found together with the amino acid sequences, respectively.

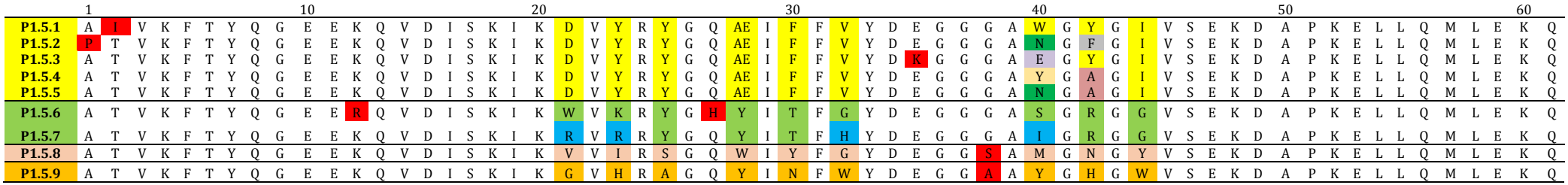


Figure 20. Amino acid sequences of nine selected clones grouped into four families with corresponding colour coding (family I, family II, family III and family IV). Single point mutations outside the binding site are indicated in red. The amino acids of the binding site are at positions: 21, 23, 25, 28, 30, 32, 40, 42 and 44 and are separated by two loops, a smaller and a larger one, at positions 26 - 27 and 33 – 39 respectively. As shown in yellow for family I (P1.5.1 – P1.5.5), an insertion occurred at position 28 within the binding site (AE, alanine+glutamate). Single point mutations outside the binding site occurred in clone P1.5.1 and P1.5.2 at the beginning of the sequence (positions 2 and 1, respectively) and in clone P1.5.3 within the larger loop (E36K). Non-assignable mutations within the binding site are coloured differently. Within family II two point mutations occurred in P1.5.6 (K12R, Q27H). In clone P1.5.8 and P1.5.9 one single point mutation was detected in each clone (G38S in P1.5.8 and G38A in P1.5.9).

Table 14. Sum of aromatic amino acids within the binding site, number of charged amino acids in the whole protein and net charge. The theoretical net charge of the protein was calculated by subtracting the number of negatively charged amino acids (D and E) from the number of positively charged amino acids (K and R).

Family	Binder	Σ aromatic amino acids (binding site)	Lysine K	Arginine R	Aspartate D	Glutamate E	Net charge of binder
I	P1.5.1	5	7	1	4	7	-3
	P1.5.2	4	7	1	4	7	-3
	P1.5.3	4	8	1	4	7	-2
	P1.5.4	4	7	1	4	7	-3
	P1.5.5	3	7	1	4	7	-3
II	P1.5.6	3	7	3	3	6	1
	P1.5.7	2	7	4	3	6	2
III	P1.5.8	3	7	1	3	6	-1
IV	P1.5.9	4	7	1	3	6	-1

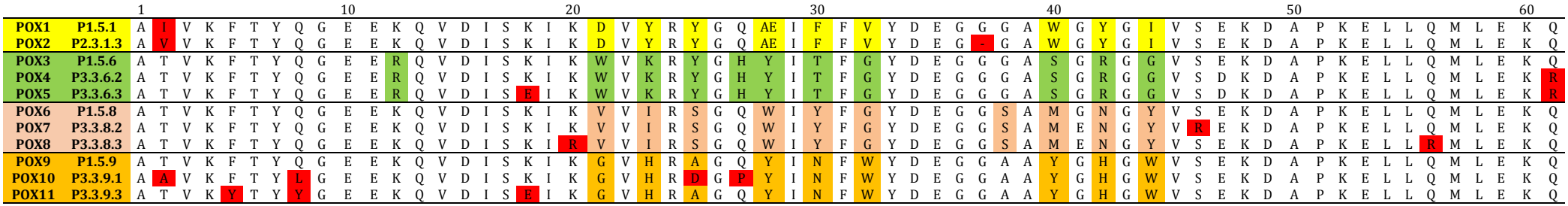


Figure 21. Table of the 11 selected Sso7d-derived peroxidasin binders. The parental clones (P1.5.1, P1.5.6, P1.5.8 and P1.5.9) are grouped together with their affinity-matured versions in four families with a corresponding colour coding as already shown in Figure 24. The final binders were renamed (e.g. P2.3.1.3 to POX2). In POX2 one point mutation (I2V) and one deletion at position 37 occurred with respect to its parental clone P1.5.1. In family II one point mutation was detected at the last position of the protein in binders POX4 and POX5 (Q61R) as well as one amino acid exchange at position 18 (K18E) in POX5 only with respect to the parental clone P1.5.6. In family III single point mutations occurred in POX7 (S46R) and POX8 (K20R, Q56R) with respect to P1.5.8. Affinity maturation of P1.5.9 yielded amino acid exchanges within the binding site (A25D) and in the small loop (Q27P) as well as at positions 2 and 8 (T2A, Q8L) in the binder POX10 and POX11 showed three mutations at positions 5, 8 and 18 (F5Y, Q8Y, K18E).

Table 15. Number of aromatic amino acids and theoretical net charge of 11 selected peroxidasin binders. The sum of aromatic amino acids within the binding site is given as well as the total number of charged amino acids. The nomenclature was changed from P1.5.1 to POX1 for reasons of readability.

Family	Binder	Σ aromatic amino acids (binding site)	Lysine K	Arginine R	Aspartate D	Glutamate E	Net charge of binder
I	POX1	5	7	1	4	7	-3
	POX2	5	7	1	4	6	-2
II	POX3	3	7	3	3	6	1
	POX4	3	7	4	4	5	2
	POX5	3	6	4	4	6	0
III	POX6	3	7	1	3	6	-1
	POX7	3	7	2	3	7	-1
	POX8	3	6	3	3	7	-1
IV	POX9	4	7	1	3	6	-1
	POX10	4	7	1	4	6	-2
	POX11	4	6	1	3	7	-3

4.3.2. Size Exclusion Chromatography

The eleven clones shown in Figure 21 were subsequently sub-cloned and expressed solubly in *E. coli* as SUMO-fusions. Next, the SUMO-tag was cleaved off with the SUMO-protease, resulting in tag-free peroxidasin binders. Of all those eleven binders a SEC analysis was performed and the elution profiles were compared to each other. All 11 binders elute, with variation in retention time, as single peaks from the column. POX 1, POX 2 and POX 9 elute later than most of the other binders (POX 3 – 8, POX 10). Another exception is POX 11. This binder elutes much later from the column. For binders POX 1 to POX 10 smaller differences in the retention time can be seen. Binders POX 1 - 10 elute between 26 to 29 minutes whereas POX 11 elutes after 43 minutes (Figure 22). No significant amount of aggregates was observed for any of the binders.

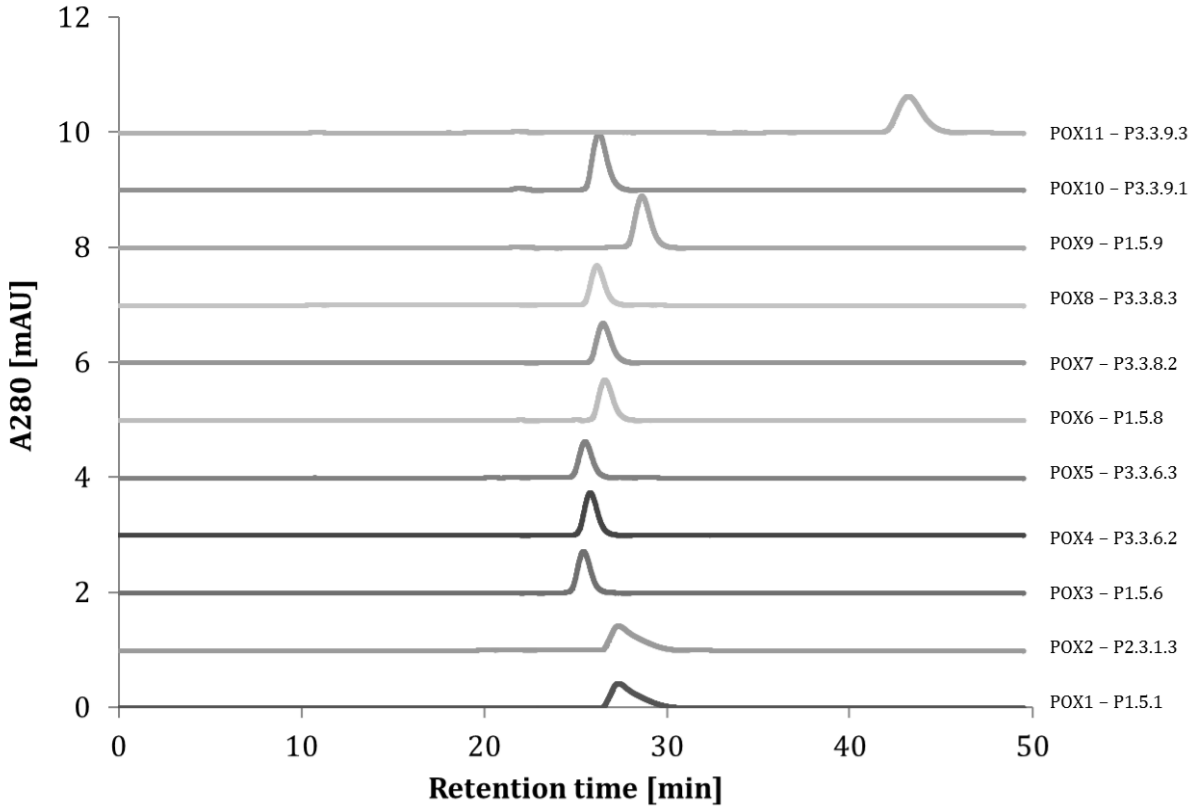


Figure 22. Size exclusion chromatography profiles of all 11 selected peroxidasin binders. 25 µg of protein were injected onto the column and detected with an UV-detector at 280 nm.

4.3.3. Analysis of Binding to Different Peroxidasin Constructs

In order to analyze the interaction of the binders with different peroxidasin constructs, the rcSso7d-mutants obtained after the 4th FACS selection round were displayed on the surface of yeast and tested for binding to 200 nM CON3, CON4, CON5 and full-length peroxidasin.

Clone P1.5.7 showed no binding signal at all. The binding signal of all tested antigens was in the range of the signal of the secondary fluorescent detection reagents only (Figure 23). In all the other 8 clones binding to the non-biotinylated version of CON4 was detected. However, binding to CON3 (leucine-rich-repeat, four Ig-like domains and heme-containing peroxidase domain) as well as full-length peroxidasin was reduced when compared with the signal obtained with CON4. Clones P1.5.6 and P1.5.8 showed no binding signal against the full-length version of peroxidasin. The highest signal was obtained with clone P1.5.8 against CON4. Together, these data demonstrate that at least most of the binders, which were engineered for binding to the peroxidasin fragment CON4, also recognize the full-length protein.

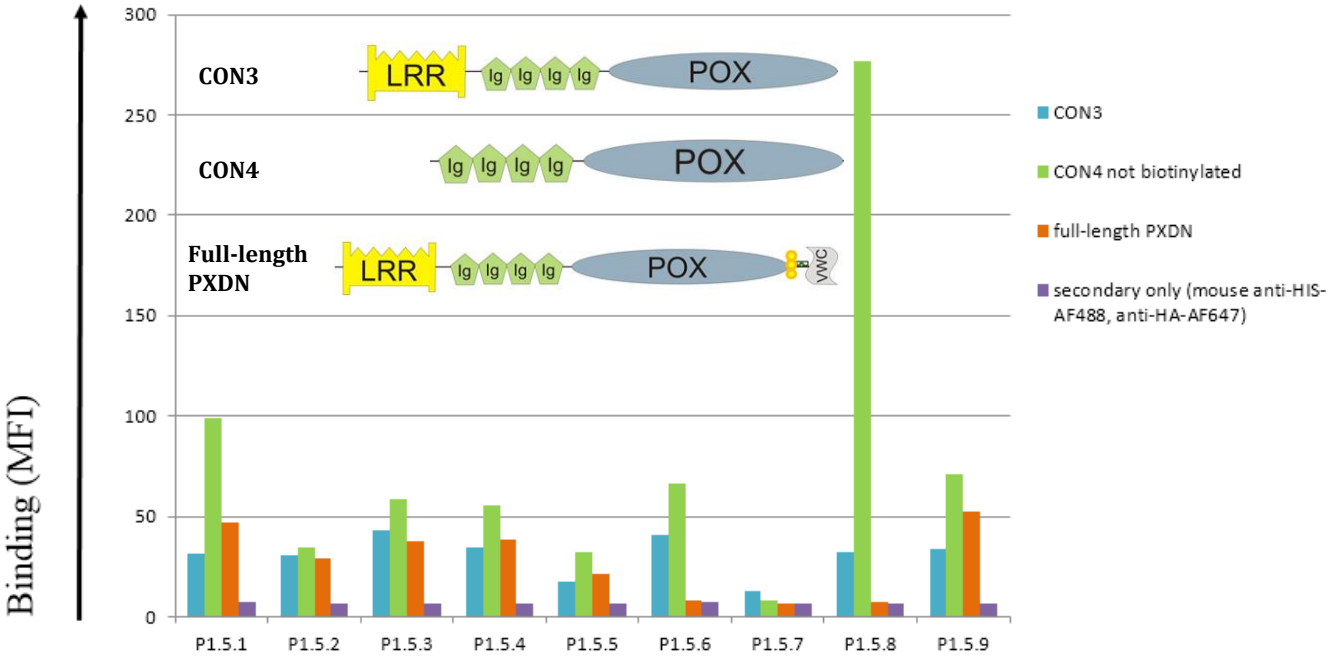


Figure 23. Measurement of binding of each clone against different antigens (CON3, CON4, full-length peroxidasin). All antigens were used at a concentration of 200 nM. All antigens contained a His-tag, which facilitated detection with the same secondary reagent (anti-HIS-AF488), thus enabling direct comparison of signal intensities between the different antigens. The three graphical representations of the peroxidasin constructs in this figure were adapted from: Soudi et al., 2015.

The data of binding to CON5, the peroxidasin variant containing the peroxidase domain only, are shown in Figure 24 below. The highest signals were obtained from clones P1.5.1 and P1.5.8. For clone P1.5.7 binding signal of CON5 was in the range of the detection signal of the secondary fluorescent detection reagents. The data of this experiment indicate that most of the binders not only have the ability to recognize CON4 (containing the peroxidase- as well as the Ig-domains) but recognize also CON5 (only containing the peroxidase domain). However, since the binding signal to CON5 is rather low in relation to the signal with the secondary reagents (Figure 24), further experiments would be needed to confirm direct binding to the peroxidase domain.

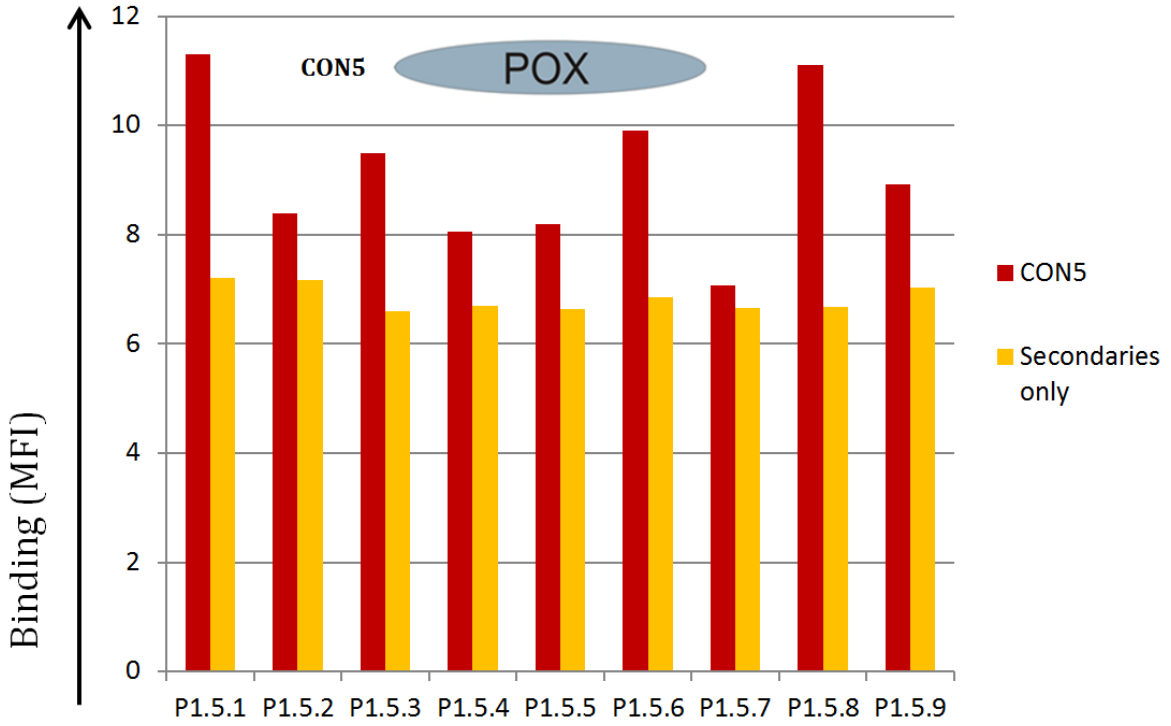


Figure 24. Measurement of binding of each clone against CON5, the peroxidasin construct containing the peroxidase domain only. CON5 was used at a concentration of 200 nM. All clones show a binding signal towards CON5 with the exception of clone P1.5.7. The binding signal of this clone is very low and slightly above the signal of the secondary reagents. The graphical representation of the peroxidasin construct in this figure was adapted from: Soudi et al., 2015.

In another experiment the parental clones together with their affinity-matured versions were measured for binding to full-length peroxidasin (200 nM). The results of this experiment are shown in Figure 25. As can be seen from this figure, the binding signal of all clones is increased in comparison to their parental clones. The only exception is clone P2.3.1.3, where the signal is slightly decreased compared to its parental clone P1.5.1. These data clearly show that the affinity-maturation process for the construction of high affinity peroxidasin binders (based on the small DNA-binding molecule Sso7d) worked. The data also show that the generated binders recognize not only the truncated versions (CON4, CON5) but also full-length peroxidasin.

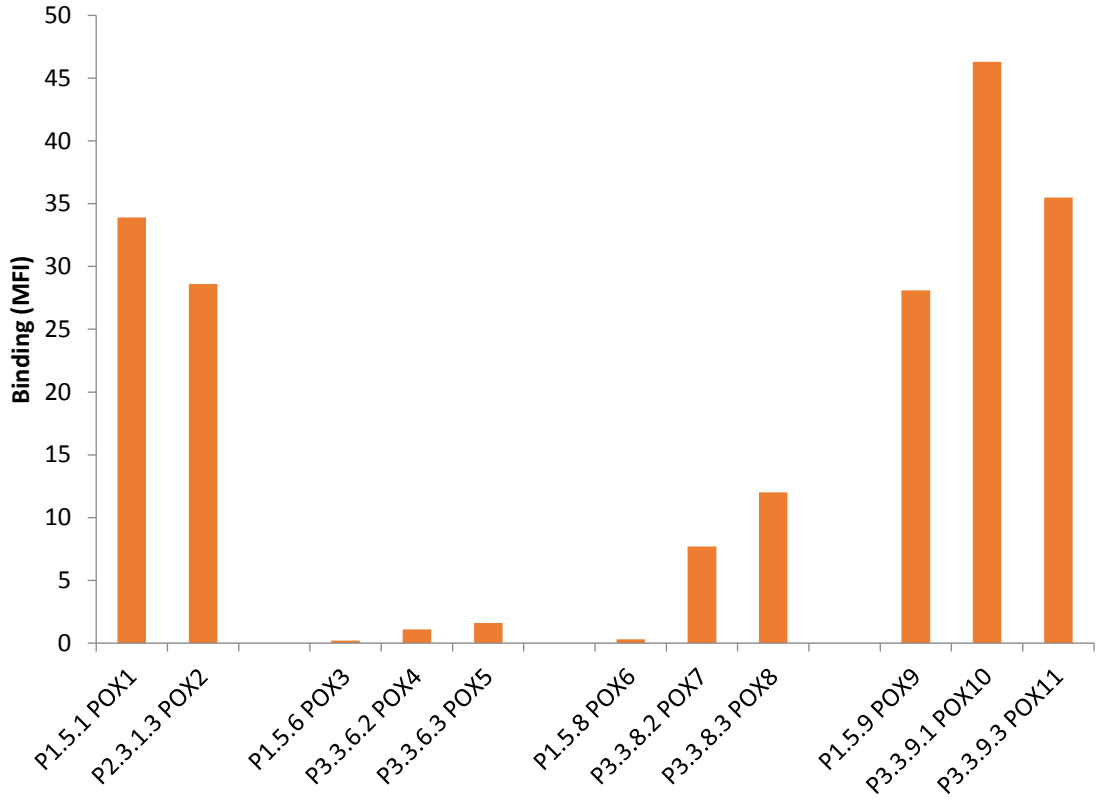
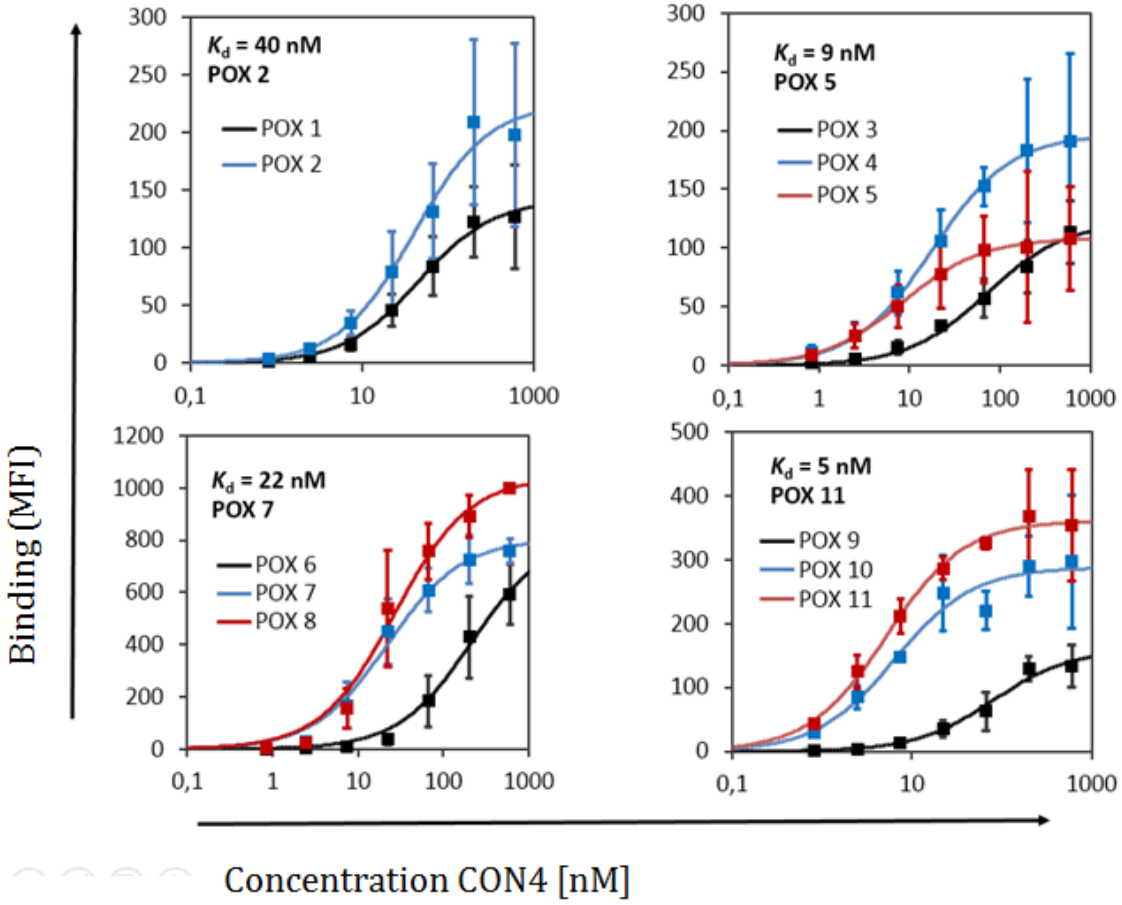


Figure 25. Measurement of binding affinity against full-length peroxidasin. Comparison of parental clones vs. affinity-maturated versions. Peroxidasin was used with a concentration of 200 nM.

Next, with all eleven selected clones titration experiments were carried out in order to determine their K_D values. The corresponding titration curves and K_D values for the parental and the affinity-matured versions of four clones are shown in Figure 26. From this figure, it can be seen clearly that the all affinity-maturated binders show a higher affinity towards CON4. The highest affinity towards CON4 was seen with binders POX10 and POX11. In comparison to their parental version their affinity has increased by more than 10-fold. The lowest increase in affinity towards the antigen CON4 can be seen between parental binder POX1 and POX2 (5 nM increase).



Binder	Clone name	K_D value [nM]
POX1 (parental)	P1.5.1	45.7
POX2	P2.3.1.3	40.4
POX3 (parental)	P1.5.6	71.2
POX4	P3.3.6.2	17.9
POX5	P3.3.6.3	8.5
POX6 (parental)	P1.5.8	217.2
POX7	P3.3.8.2	22.1
POX8	P3.3.8.3	26.9
POX9 (parental)	P1.5.9	80.8
POX10	P3.3.9.1	6.4
POX11	P3.3.9.3	5.2

Figure 26. Titration of parental and affinity-maturated versions of peroxidasin binders plus the results from the determination of the K_D values of eleven selected clones. All experiments were executed in triplicates and a mean value was calculated. On each chart only one K_D value of an affinity-matured binder is shown (best binder of each family is shown e.g. on the bottom right corner POX 11 with a K_D of 5 nM). The black curves indicate the wild-type clones, blue and red lines are showing the mutated and affinity-matured versions. The 8-point curves were obtained in three independent experiments. Data were fitted by using a 1:1 binding model and the resulting fitted curves and K_D values are shown.

4.3.4. Competition Experiments

Competition experiments were executed in order to test if the obtained binders recognize the same or different epitopes on the surface of peroxidasin (CON4).

The competition assay was performed in duplicates. From each of the four binder families the best binder was selected for the experimental set up (POX2, POX5, POX7 and POX11). The results of both experiments are shown in Figure 27, demonstrating that the results are highly reproducible. The grey colour code shows the residual binding in percent after blocking. The darker the grey value the less binding occurs after blocking. If a yeast cell expressed on its surface an rcSso7d-mutant that competes for CON4-binding with the binder which was used in the preincubation step with CON4, no signal (or a strongly reduced signal) was obtained due to the fact that the binders recognized the same (or an overlapping) epitope. One can clearly see this result in the diagonal line of the graphic (black, no residual binding), conforming that all binders block themselves. A different result can be seen for certain combinations of binders derived from different sequence families. Between binder POX2 and POX5 a residual binding strength of around 70 - 90 % occurred in both experiments suggesting that those two binders recognize different epitopes on peroxidasin. The combination of POX2 with POX 7 and POX11 respectively led to no residual binding, suggesting that those binders have overlapping epitopes. Binder POX5 expressed on the yeast surface combined with POX2 and POX11 showed stronger residual binding after blocking compared to POX5 with POX7. Interestingly enough when POX7 was expressed on the cell surface and mixed with POX5 no signal was detected. Expressed POX11 on the surface in combination with POX5 led to residual binding of around 80 to 90 %. Together, these data indicate that the engineered binders recognize at least two different epitopes on CON4. More specifically, POX5 most probably binds to an epitope that is different from those of the three other binders.

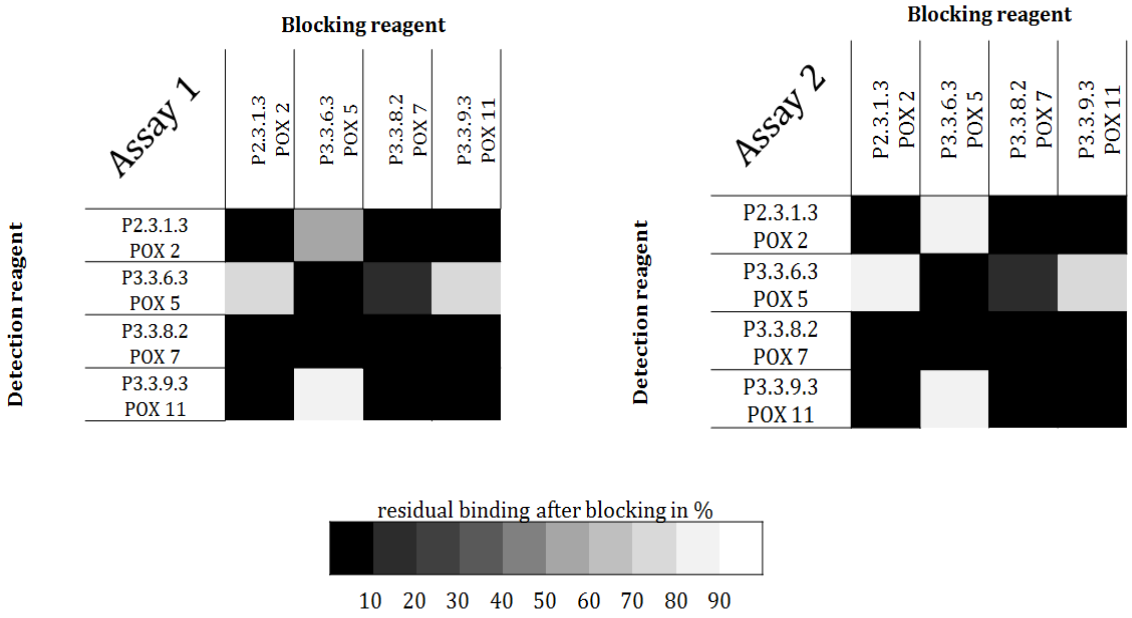


Figure 27. Competition experiment executed with four different binders derived from four different sequence families. The assay was performed in two different experiments (left and right heat map, respectively). The left column (detection reagent) is the binder expressed on the yeast surface. The blocking reagent was the solubly expressed binder pre-incubated with antigen CON4. The grey-scale below the two heat maps shows the residual binding after blocking (in %), compared with the signal obtained in the absence of any blocking reagent.

4.3.5. Differential Scanning Calorimetry (DSC)

The thermal stability of all binders and antigen CON4 was determined by differential scanning calorimetry. Furthermore, combinations of the antigen CON4 together with selected binders were used to test if the binder can stabilize the antigen, leading to a rise in T_m values. Due to problems during the experiment with the DSC-device (VP-Capillary DSC System from MicroCal) not all combinations could be measured successfully.

The transition temperatures of the POX binders (35 µM) varied between 60.1°C and 93.8°C (Table 16), demonstrating that all selected mutants are thermally stable proteins. There are two transition temperatures for the peroxidasin construct CON4: T_m1 and T_m2 of 50.4°C and 74.6°C respectively (Table 17). Combining CON4 with either POX1 or POX2 led to a rise of 1.5°C for the first transition temperature and 1.0°C (POX1) and 2.1°C (POX2) for the second transition temperature (Figures 28A and 28B). Binder POX3 seemed to have no stabilizing effect on CON4 (same T_m values for both transitions like in CON4 only). POX7 and POX8 addition to CON4 led to a rise in temperature only for the first transition. T_m values for the second transition even dropped down to 74.0°C (CON4 + POX7) and 73.9°C (CON4 + POX8). POX10 showed the strongest stabilizing effect on CON4, increasing the T_m values from 50.4 to 52.9°C and from 74.6 to 76.5°C. In the two tables below (Table 16 and Table 17) all obtained results are shown.

Table 16. Results from the differential scanning calorimetry experiment for all eleven peroxidasin binders (35 µM). The measurements were carried out in two different experiments. Due to problems during the measurement not all data could be obtained. Missing values are indicated by n.a. (not available).

Family	Sample (35 µM)	T_m [°C] 1st measurement	T_m [°C] 2nd measurement	T_m [°C] average
I	POX 1	93.8	n.a.	93.8
	POX 2	n.a.	n.a.	n.a.
II	POX 3	81.4	78.4	79.9
	POX 4	72.0	71.9	72.0
	POX 5	63.7	63.4	63.6
III	POX 6	n.a.	n.a.	n.a.
	POX 7	70.8	71.2	71.0
	POX 8	n.a.	n.a.	n.a.
IV	POX 9	80.4	n.a.	80.4
	POX 10	80.7	79.6	80.2
	POX 11	n.a.	60.1	60.1

Table 17. Results of the differential scanning calorimetry experiment measuring the antigen CON4 without added POX binders and in combination with various binders. In the setup binder + antigen (CON4) the used concentrations were 10 µM for the binder and 5 µM for CON4, respectively.

Sample	T_m1 [°C]	T_m2 [°C]
CON4 only	50.4	74.6
CON4 + POX 1	51.9	75.6
CON4 + POX 2	51.9	76.7
CON4 + POX 3	50.7	74.3
CON4 + POX 7	52.8	74.0
CON4 + POX 8	54.0	73.9
CON4 + POX 10	52.9	76.5

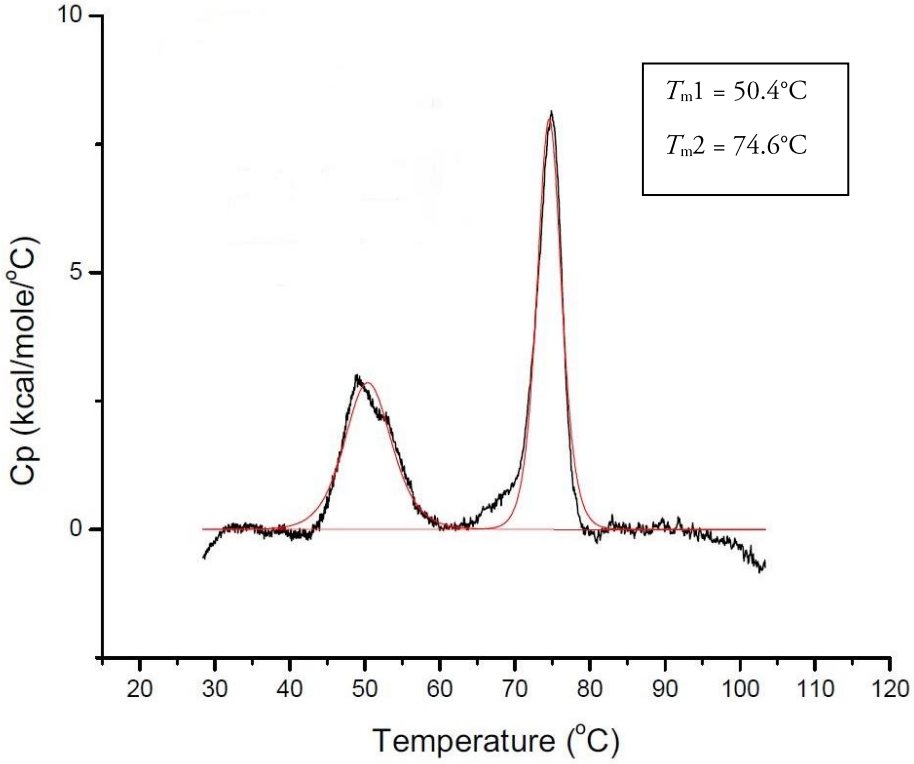


Figure 28A. DSC profile of peroxidasin construct CON4. The black line represents the raw data whereas the fitted curves are shown in red. The thermogram shows two peaks. The midpoints of the two peaks represent the corresponding temperature at which half of the respective domain is denatured (peroxidase and IgG-like domain).

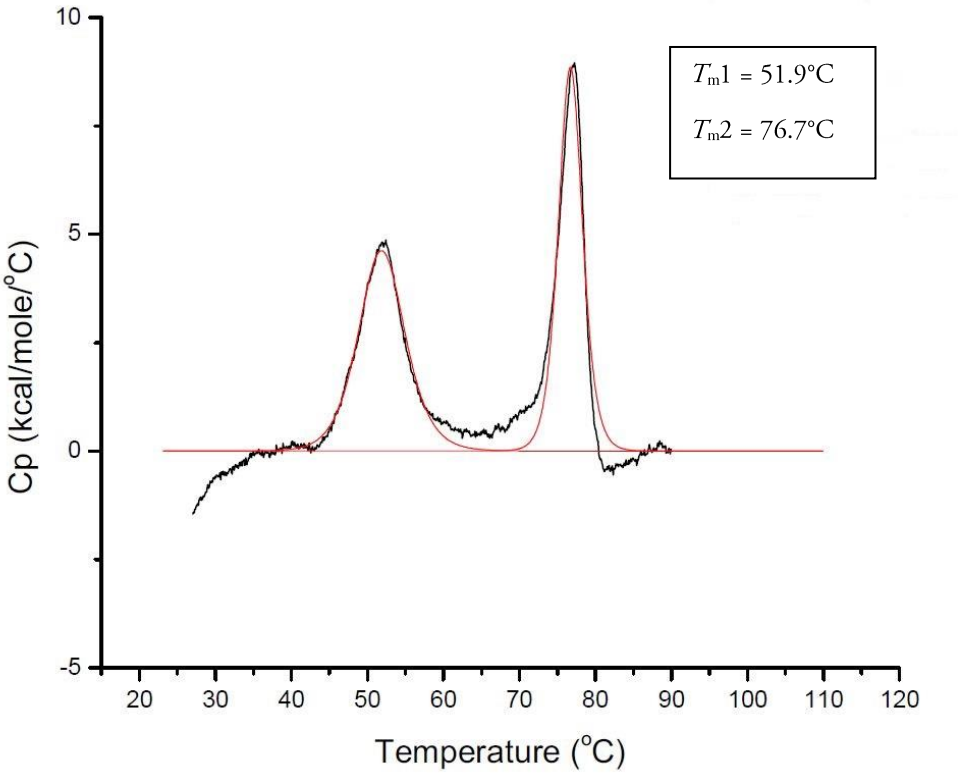


Figure 28B. DSC profile of CON4 stabilized with POX2.

5. Discussion

During this study binders were engineered against truncated versions of human peroxidasin as well as full-length peroxidasin using two yeast surface display libraries (rcSso7d-11, rcSso7d-18) of a charge-reduced version of the DNA-binding protein Sso7d from the archaeon *Sulfolobus solfataricus* (rcSso7d). From the already existing yeast libraries binders were selected through several rounds of affinity maturation using magnetic bead selection and FACS. Finally, we ended up with 11 binders with affinities (K_D values) ranging from 200 nM down to 5 nM, the latter resembling already the affinity of a typical IgG-antibody. The future use of these rcSso7d-derived peroxidasin binders could include the support of the full-length peroxidasin purification process. Furthermore, the binders could be used as chaperones for crystallization experiments of peroxidasin and its truncated versions. Another potential application would be the purification of endogenous peroxidasin from blood plasma e.g. via affinity chromatography. Finally peroxidasin binders could be used as detection reagent, e.g. in an enzyme-linked immunosorbent assay (ELISA).

In summary it can be stated, that the small reduced-charge variant of the DNA-binding protein Sso7d can be effectively engineered for molecular recognition of different epitopes on human peroxidasin and with affinities down to the single-digit nanomolar range.

5.1. Selection Process

After two rounds of magnetic bead selection (in order to retain weak affinity binders), one round of error-prone PCR and another round of a magnetic bead selection four consecutive rounds of FACS enrichments were carried out for further selection of high affinity binders. From library P1.5 the first 9 single clones were obtained and analysed. We clustered these clones into four families according to their amino acid sequence similarities. The binding site (located on the rigid triple-stranded β-sheet on the surface of the protein) of those enriched binders contained two to five aromatic amino acids. The net charge of the protein covered a spectrum from slightly positive with a net charge of +1 to a surplus of negative charges down to a net charge of -3. A mixture of clones (P1.5.1, P1.5.2, P1.5.3 and P1.5.4) as well as clones P1.5.6, P1.5.8 and P1.5.9 were selected for further rounds of mutagenesis and were subjected to further selection rounds using fluorescent activated cell sorting (FACS). All clones were selected according to the affinity of the surface-expressed binder towards our desired antigen peroxidasin and its constructs we used during this work. Another factor we considered for the selection of the clones which were subjected to further rounds of mutagenesis and sorting, was the expression level of binders (data not shown) on the surface of the yeast cells. Clones showing a higher expression rate were preferred. Finally, we sequenced 17 affinity matured clones from which we chose seven. The initial four parental clones (P1.5.1, P1.5.6, P1.5.8 and P1.5.9) together with the seven clones led to our final 11 clones.

During the fluorescence activated cell sorts a constant shift towards higher affinity could be determined. That implies that clones with higher affinity towards the antigen could be successfully enriched. This finding is also reflected by the improved K_D values of the affinity matured mutants when compared with their parental counterparts. This demonstrates that the introduction of mutations by ep-PCR and subsequent selection for improved binding led to beneficial amino acid variations within the protein.

5.2. Binder characterization

The above mentioned eleven binders (4 parental clones and 7 affinity matured versions) were used for further characterization including yeast surface titration in order to determine their affinity (K_D value). Furthermore, competition experiments on CON4 were executed in order to investigate whether the engineered binders recognize different and/or overlapping epitopes.

Biophysical characterization included size exclusion chromatography (SEC) and differential scanning calorimetry (DSC).

As shown in Figure 22 the binders eluted from the column at different points in time. Ten out of eleven binders elute between 25 - 30 minutes. Binder POX11 was eluted from the column after approximately 43 min, most probably due to interaction with the column material. This binder contains four aromatic amino acids within the binding site. The net charge of POX11 is -3. Other binders e.g. those from Family I (POX1, POX2) contain even five aromatic amino acids within their binding site. The unique feature of POX11 is the gain of two tyrosines at positions 5 and 8. These two tyrosines could explain the stickiness and longer elution time of the protein. All eleven binders eluted as single peaks suggesting that they are monomeric. No aggregation was observed.

Analysis of the thermal stability by DCS revealed that most of the eleven binders show T_m values above 70°C except for POX5 (63.6°C) and POX11 (60.1°C). Adding POX1, POX2 and POX10 to CON4 led to a clear rise in transition temperature indicating that these two binders have a stabilizing effect on CON4. Therefore, binders POX1, POX2 and POX10 could function as crystallization chaperones in order to enhance the stability during the process of crystal formation.

From the four families of binders the affinity-matured versions of the parental clone POX9 (POX10 and POX11) showed the lowest K_D values. Both bind in the single-digit nanomolar range to the peroxidasin construct CON4 (K_D values of 6.4 and 5.2 nM respectively). This represents a more than 10-fold increase in affinity in comparison with their parental clone. The least improvement could be seen in Family I comprising only two binders (POX1, POX2). The difference between parental and affinity matured binder is one point mutation (I2V) and one deletion at position 37. The similarity between the amino acid sequences of the two proteins may explain the lack of affinity improvement towards CON4 (45.7 nM for POX1 versus 40.4 nM for POX2). In Family III a 10-fold increase in affinity was measured (POX6: 217.2 nM, POX7: 22.1 nM and POX8: 26.9 nM). From Family II one binder showed a binding affinity in the single-digit nanomolar range (POX5). Summarising, the highest affinity against CON4 was obtained with clones POX5, POX10 and POX11. All three show K_D values in the single-digit nanomolar range.

With the best binder of each family (POX2, POX5, POX7 and POX11) competition experiments were executed in order to gain more information about the epitope to which the proteins bind on the antigen. Combining binders POX2 + POX5 and POX5 + POX11 showed residual binding after blocking which strongly suggests that POX5 recognizes an epitope that is different from those of clones POX2 and POX11. A weaker binding after

blocking can be seen in the combination of binders POX5 (on yeast cell surface) + POX7 (residual binding of 20 – 30 %) indicating that POX5 and POX7 may partly block binding of each other through allosteric effects. Interestingly, mixing yeast cells expressing POX7 on the surface together with soluble expressed POX5 led to no residual binding after blocking.

It could also be shown that the binders not only recognize the truncated versions of peroxidasin (CON4-B and CON5) but also the full-length version.

6. Bibliography

Ackerman, M., Levary, D., Tobon, G., Hackel, B., Orcutt, K. D., & Wittrup, K. D. (2009). Highly avid magnetic bead capture: an efficient selection method for de novo protein engineering utilizing yeast surface display. *Biotechnology progress*, *25*(3), 774-783.

Angelini, A., Chen, T. F., de Picciotto, S., Yang, N. J., Tzeng, A., Santos, M. S., ... & Wittrup, K. D. (2015). Protein engineering and selection using yeast surface display. In *Yeast Surface Display* (pp. 3-36). Humana Press, New York, NY.

Aylon, Y., & Kupiec, M. (2004). New insights into the mechanism of homologous recombination in yeast. *Mutation Research/Reviews in Mutation Research*, *566*(3), 231-248.

Baumann, H., Knapp, S., Lundbäck, T., Ladenstein, R., & Härd, T. (1994). Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus solfataricus*. *Nature Structural and Molecular Biology*, 1(11), 808-819.

Bhave, G., Cummings, C. F., Vanacore, R. M., Kumagai-Cresse, C., Ero-Tolliver, I. A., Rafi, M., ... & Hudson, B. G. (2012). Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis. *Nature chemical biology*, *8*(9), 784-790.

Boder, E. T., & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology*, *15*(6), 553-557.

Gera, N., Hussain, M., & Rao, B. M. (2013). Protein selection using yeast surface display. *Methods*, *60*(1), 15-26.

Khoshnoodi, J., Pedchenko, V., & Hudson, B. G. (2008). Mammalian collagen IV. *Microscopy research and technique*, *71*(5), 357-370.

Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., & Ladenstein, R. (1996). Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophilesulfolobus solfataricus. *Journal of molecular biology*, *264*(5), 1132-1144.

Lázár, E., Péterfi, Z., Sirokmány, G., Kovács, H. A., Klement, E., Medzihradszky, K. F., & Geiszt, M. (2015). Structure–function analysis of peroxidasin provides insight into the mechanism of collagen IV crosslinking. *Free Radical Biology and Medicine*, *83*, 273-282.

McCall, A. S., Cummings, C. F., Bhave, G., Vanacore, R., Page-McCaw, A., & Hudson, B. G. (2014). Bromine is an essential trace element for assembly of collagen IV scaffolds in tissue development and architecture. *Cell*, *157*(6), 1380-1392.

Nelson, R. E., Fessler, L. I., Takagi, Y., Blumberg, B., Keene, D. R., Olson, P. F., ... & Fessler, J. H. (1994). Peroxidasin: a novel enzyme-matrix protein of Drosophila development. *The EMBO Journal*, *13*(15), 3438-3447.

Paumann-Page, M., Katz, R. S., Bellei, M., Schwartz, I., Edenhofer, E., Sevcnikar, B., ... & Obinger, C. (2017). Pre-steady-state kinetics reveal the substrate specificity and mechanism of halide oxidation of truncated human peroxidasin 1. *Journal of Biological Chemistry*, 292(11), 4583–4592.

Péterfi, Z., & Geiszt, M. (2014). Peroxidasins: novel players in tissue genesis. *Trends in biochemical sciences*, *39*(7), 305-307.

Soudi, M., Zamocky, M., Jakopitsch, C., Furtmüller, P. G., & Obinger, C. (2012). Molecular evolution, structure, and function of peroxidasins. *Chemistry & biodiversity*, *9*(9), 1776-1793.

Soudi, M., Paumann-Page, M., Delporte, C., Pirker, K. F., Bellei, M., Edenhofer, E., ... & Van Antwerpen, P. (2015). Multidomain human peroxidasin 1 is a highly glycosylated and stable homotrimeric high-spin ferric peroxidase. *Journal of Biological Chemistry*, 290(17), 10876-90.

Traxlmayr, M. W., Kiefer, J. D., Srinivas, R. R., Lobner, E., Tisdale, A. W., Mehta, N. K., ... & Wittrup, K. D. (2016). Strong enrichment of aromatic residues in binding sites from a charge-neutralized hyperthermostable Sso7d scaffold library. *Journal of Biological Chemistry*, 291(43), 22496-22508.

Vanacore, R., Ham, A. J. L., Voehler, M., Sanders, C. R., Conrads, T. P., Veenstra, T. D., ... & Hudson, B. G. (2009). A sulfilimine bond identified in collagen IV. *Science*, *325*(5945), 1230-1234.

Weiss, S. J. (2012). Peroxidasin: tying the collagen-sulfilimine knot. *Nature chemical biology*, *8*(9), 740-741.

Zederbauer, M., Furtmüller, P. G., Brogioni, S., Jakopitsch, C., Smulevich, G., & Obinger, C. (2007). Heme to protein linkages in mammalian peroxidases: impact on spectroscopic, redox and catalytic properties. *Natural product reports*, *24*(3), 571-584.

7. List of Figures

8. List of Tables