Functional insights on ZPLD1 (Cupulin)

- a structural protein of the equilibrium organ -

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Marie Isabell Knepper

2023

This work was prepared between the 1st of April 2020 and the 31st of March 2023

with guidance of Dr. Jens Dernedde

at the Institute of Laboratory Medicine, Clinical Chemistry and Pathobiochemistry

of Charité – Universitätsmedizin Berlin

1st reviewer: Dr. Jens Dernedde

Charité – Universitätsmedizin Berlin, Institute of Laboratory Medicine, Clinical Chemistry and Pathobiochemistry

2nd reviewer: Prof. Dr. Rainer Haag

Freie Universität Berlin, Institute of Chemistry and Biochemistry

Date of defense: 12.01.2024

Acknowledgement

From the bottom of my heart, I would like to thank all, who supported me in this journey towards my finished doctoral thesis. Be it with expertise, guidance, emotional backing, or resources such as time and money. Special thanks to:

- **Jens** for being the best supervisor I could have wished for. For providing me with the opportunity to step back into research in the first place and for all your support on the way. For your calmness, smartness and all the fun we had.
- **Burak, Marten, Julian, and Luise** for being such a great team. Making it fun to come to the lab even when results have at times been unsatisfying. Thank you for having open ears, to ask any stupid question one might think of, and for always finding solutions together.
- **Kai** for supporting me with the EM experiments. I still think those are super exciting and I am sad that we were not able to dig deeper into them.
- **Prof. Dr. Rainer Haag** for the kindness to become my second supervisor without even knowing me.
- **Jonas, my (extended) Family and my friends**, who believed in me all the time, without any restrictions. Who did not call me crazy for spending more than three years of my life researching a little protein from the inner ear. Thank you for your emotional as well as practical support throughout this time.
- **BCG**, especially my career advisors **Sebastian and Constantin** as well as my office and line manager **Torsten**, who approved my rather untypically request for an educational leave. Furthermore, the firm for providing me with a two-year scholarship as well as the opportunity to work part-time to enable me financing this period of my life.
- **Charité Berlin** and the Institute of Laboratory Medicine, Clinical Chemistry and Pathobiochemistry including all employees who have crossed my way there, for providing the environment and save working place to pursue my research.

I am grateful for all the time, dedication and trust invested in me. Thank you so much.

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me.

This dissertation has not been accepted or rejected in any other previous doctorate examination.

Marie Isabell Knepper

Berlin, 28.07.2023

Table of content

Abstract

Zona Pellucida Like Domain 1 Protein (ZPLD1) is a main component of the cupula, a gelatinous structure located in the labyrinth organ of the inner ear and involved in vestibular function. The N-glycosylated protein is likely able to organize high-molecular-weight polymers via its zona pellucida (ZP) module, which is common for many extracellular proteins that self-assemble into matrices. In this work, I confirmed that ZPLD1 can form multimers while setting up a cellular model leveraging Mardin-Darby canine kidney (MDCK) cells to study protein polymerization. I identified two motifs within ZPLD1 which regulate protein polymerization and follow previously published conserved regions, identified across ZP proteins. Mutational depletion of either one of these modules led to diminished or abnormal polymer formation outside of the cells, likely due to altered processing at the plasma membrane. Further, intracellular polymer formation was observed. Proteolytic cleavage during secretion, separating the regulatory motif located distinct of the ZP module from the mature monomer, seems to be necessary to enable proper polymerization. This separation requires a transmembrane domain as well as the consensus cleavage site, which were as well located and confirmed using the established model and corresponding ZPLD1 expression variants. I was in addition able to proof that a serine type protease, potentially membrane bound hepsin, executes protein maturation. While the molecular interactions of the identified motifs and the exact processing protease and mechanism remain to be proven, my findings suggest that ZPLD1 is a polymer forming ZP protein, following a well-orchestrated mechanism. This mechanism is likely shared by other homopolymerizing ZP family proteins, includes cleavage of a pro-peptide, conformational changes, and the formation of precursor protein dimers. Finally, the mechanism leads to building up the gelatinous hydrogel of the cupula, thereby being of high significance for proper vestibular function. Ultimately, the variety of pathologies with reported impact of altered ZPLD1 expression indicate potential further functionalities such as cell adhesion, migration or segregation adding significance of this protein as future research target.

Zusammenfassung

Zona Pellucida Like Domain 1 Protein (ZPLD1) ist ein Hauptbestandteil der Cupula. Die Cupula ist eine gelartige Struktur, die sich im Labyrinth Organ des Innenohres befindet und für die räumliche Orientierung von entscheidender Bedeutung ist. Das N-glycosylierte Protein ist sehr wahrscheinlich in der Lage sich über sein Zona Pellucida (ZP) Modul zu komplexen, höhermolekularen Polymeren zusammenzuschließen. Diese ZP-Module sind typischer Bestandteil zahlreicher polymerbildender Proteine aus der Familie der ZP-Proteine. In der vorliegenden Arbeit konnte ich mithilfe eines zellulären Modells nachweisen, dass ZPLD1 Multimere bilden kann. Das Modell basiert auf stabil transfizierten Mardin-Darby-canine kidney (MDCK) Zellen, mit denen ich verschiedener ZPLD1 Varianten untersuchen und erste Erkenntnisse zu den zugrundeliegenden Mechanismen und Voraussetzungen gewinnen konnte. Dabei habe ich zwei hydrophobe Regionen identifiziert, die bei ZPLD1 wie auch anderen ZP-Proteinen eine wichtige regulatorische Funktion übernehmen. Deletion dieser Sequenzregionen führt zu abnormaler oder gar vollständig reduzierter Polymerbildung. Die beruht mutmaßlich auf einer fehlerhaften Prozessierung an der Zellmembran während des Sekretionsprozesses. Für diese Varianten konnten außerdem intrazelluläre Polymere beobachtet werden. Proteolytische Abspaltung während der Sekretion und die damit verbundene Trennung der regulatorischen Motive scheint unabdingbar für die korrekte Polymerbildung von ZPLD1 zu sein. Diese Spaltung bedarf zunächst einer Membranverankerung durch eine Transmembrandomäne sowie der Existenz einer Consensus Spaltstelle, die ich ebenfalls mittels oben genannter Methoden identifizieren und nachweisen konnte. Ich konnte zeigen, dass eine Serin-Protease für die Prozessierung von ZPLD1 verantwortlich ist. Obwohl der genaue Mechanismus und die exakte Protease noch identifiziert und beschrieben werden müssen, deuten meine Ergebnisse darauf hin, dass die Polymerbildung auf einem fein abgestimmten Mechanismus beruht. Dieser Mechanismus scheint dabei weitgehend identisch mit dem anderer Homopolymer bildender ZP-Proteine zu sein und bedarf der Spaltung des Pro-Peptides, Konformationsänderungen sowie die Bildung dimerer Vorläuferpolymere. Dieser Prozess ermöglicht die Ausbildung der Cupula und ist damit von entscheidender Bedeutung

für die räumliche Orientierung. Veröffentlichte Transkriptom-Daten verschiedener Pathologien lassen überdies vermuten, dass ZPLD1 in Teilen wichtig für Prozesse wie Zell-Zell-Interaktion und -Adhesion, Zell-Migration oder die zellulären Segration während der Embryonalentwicklung sein könnte. Daher stellt ZPLD1 weiterhin ein spannendes Ziel für zukünftige Forschung dar.

1 Introduction

1.1 ZPLD1 origin and localization within the labyrinth organ of the inner ear

Zona Pellucida Like Domain Protein 1 (ZPLD1), also known as Cupulin, is a 45 kDa glycoprotein initially identified as a major part of the cupula, located in the labyrinth organ of the inner ear (Figure 1) (Dernedde et al., 2014).

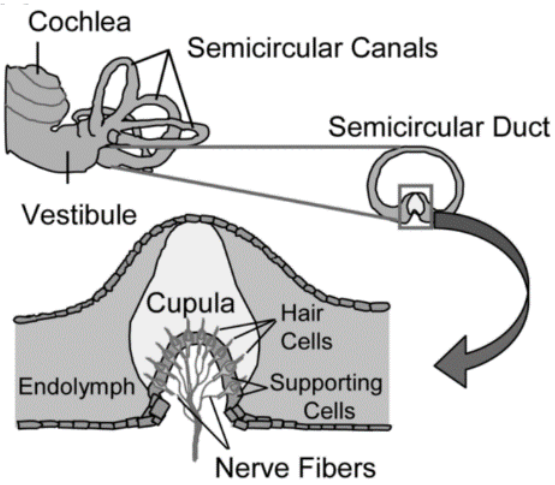


Figure 1. Schematic representation of the localization of the cupula in the (human) inner ear (Dernedde et al., 2014). Figure from Ref. 1 (Dernedde et al., 2014).

The cupula functions as a sensory hydrogel in the liquid filled cavity at the basis of the semicircular canals. The structure is fixed at the roof of the ampulla and rides on a barrel-like structure, the crista ampullaris. There, the gelatinous cupula is connected with the underlying neuroepithelium by kino- and stereocilia growing out from the top of hair cells. The hair cells are stimulated upon torsional acceleration of the head and corresponding rotation of the inner ear fluid. The resulting deflection of the cupula and subsequent bending of the cilia opens mechanosensitive ion channels of the hair cells enabling a potassium influx and the set-up of a generator potential (Dohlman, 1969; Helling et al., 2000; Takumida, 2001). The afferent bipolar nerve exhibits an alteration of the action potential rate and transmits the signal further towards the brain. This

stimulation mechanism is impeded once the cupula leaks or detaches from its roof (Helling et al., 2000).

Malformations in the ampulla are therefore considered as a possible explanation for the sudden loss of vestibular function. Experiments in pigeons, investigating mechanical detachment of the cupula from the roof of the ampulla support this assumption (Scherer and Watanabe, 2001).

1.2 ZPLD1 potential biological function in the cupula and beyond

To date the biological function of ZPLD1 as part of the cupula is not yet well studied. However, loss of structural integrity of the cupula might be caused by a lack of structural material production (Iimura et al., 2010). Correspondingly, spontaneous mutations in the mouse ZPLD1 gene have been shown to lead to circling behavior of mutant mice, thereby indicating balance dysfunction (Vijayakumar et al., 2019). ZPLD1-/- mice exhibit normal hearing function and their otolithic organs appear normal while only loss of sensory input for rotary movements was observed (Vijayakumar et al., 2019).

In humans, a ZPLD1 gene mutation leading to a 2.5-fold decreased expression level has been described in a patient exhibiting cerebral cavernous malformations (CCM). Interestingly, no alteration of vestibular function was reported for the patient (Gianfrancesco et al., 2008). Studies, searching for a causal context of aberrations of diastolic blood pressure reported an area containing the ZPLD1 gene as most significant chromosomal region (Liu et al., 2016). These observations are particularly interesting as they suggest that ZPLD1 might be part of a complex signaling pathway implicated in CCM and blood vessel formation and may exhibit other functions besides structuring the cupula.

In addition, several genome- and/or transcriptome-wide analyses which postulate ZPLD1 gene mutations related to childhood obesity, pancreatic cancer, sensory nerve disturbances and breast cancer metastasis suggest possible roles e.g., for cell-cell adhesion, migration and (nervous system) development (Chen et al., 2019; Glessner et al., 2010; Kobayashi et al., 2013; Meng et al., 2020). The multitude of possible tasks for

ZPLD1 thereby mirrors the diversity of ZP protein function, from acting as structural components of the mammalian egg coat (the zona pellucida) and other polymeric structures to serving as receptor, tumor suppressor or extracellular network to capture pathogens (Jovine et al., 2005; Pak et al., 2001).

1.3 The ZP protein family and their common characteristics

ZP proteins are extracellular glycoproteins containing a zona pellucida (ZP) module, a structural element consisting of ~260 amino acids with eight conserved cysteine residues (Bork and Sander, 1992). Many ZP proteins are involved in the formation of filaments and/or matrices, which is consistent with the role of the module in protein polymerization (Jovine et al., 2005). Numerous studies suggest a common role and function of the bipartite ZP module with its N-terminal and C-terminal domains (ZP-N and ZP-C) linked via an interdomain that structurally organizes the module in a distinct manner (Bokhove and Jovine, 2018). Other common features of most ZP proteins are a membrane anchoring domain, a conserved consensus cleavage site (CCS) and two hydrophobic patches (Jovine et al., 2005). These hydrophobic patches, one within the ZP module (internal hydrophobic patch / IHP) and one located distal of the ZP module (external hydrophobic patch / EHP) seem to play a particular relevant role to regulate protein assembly (Jovine et al., 2004). Cleavage of ZP precursors during secretion was proposed to lead to a loss of the regulatory EHP, thereby enabling the released mature polypeptide to assemble (Jovine et al., 2004). Structural insights based on X-ray studies and cryo-EM models of Uromodulin revealed, that the ZP module function relies on the ability to support homodimerization (Bokhove et al., 2016; Stanisich et al., 2020; Stsiapanava et al., 2020). Homodimers, resulting from an interplay between IHP and EHP of two separate molecules present a conformation which seems to be crucial for the described cleavage and might further prevent intracellular polymerization (Bokhove et al., 2016). For some ZP proteins, additional structural elements of the interdomain linker have been shown to be of high relevance and involved in maintaining this polymerization competent protein conformation (Bokhove et al., 2016). The

organization and type of the linker seems to play a significant role for polymeric assembly of these ZP proteins, enabling conformational changes and domain swapping (Bokhove et al., 2016; Han et al., 2010). These mechanisms facilitate the well-orchestrated scenario of IHP and EHP head-to-tail module assembly and the formation of polymers (Han et al., 2010; Stsiapanava et al., 2020).

1.4 Research question and aim of the presented project

Amino acid identity of ZPLD1 to classical ZP proteins is not very striking and only based on certain especially relevant conserved amino acids. E.g., overall sequence identity between zebrafish ZPLD1 and the well-studied ZP proteins (including structural details) human Uromodulin and mouse Zona pellucida sperm-binding protein 2 is less than 15% and less than 12% respectively. In contrast, interspecies identity between zebrafish, salmon and human ZPLD1 adds up to more than 70% (Figure 2).

Despite the low identity to other ZP proteins, I used common knowledge of their structure and functioning to build my hypotheses for the presented work and research. Correspondingly, I hypothesized that ZPLD1 gets cleaved by a protease at the consensus cleavage site (CCS) during secretion whilst anchored to the plasma membrane. The N-terminal signal peptide (SP) guides ZPLD1 to the vesicular transport system preparing for secretion while a transmembrane domain (TMD), located within the C-terminal pro-peptide would ensure tight membrane association and essential positioning. Cleavage at the specific CCS at the C-terminus of ZPLD1 releases the pro-peptide and thereby the interaction between two hydrophobic regions, leading to a conformational activation of the ZP module and subsequent protein polymerization. One of these hydrophobic patches is proposed to be located within the ZP module and is therefore called internal hydrophobic patch (IHP) while the other is expected to be located outside of the ZP module as part of the pro-peptide and is therefore called external hydrophobic patch (EHP) (Figure 3).

```
CLUSTAL 2.1 multiple sequence alignment


Zebrafish       MERLCVILLLVSKTFIANAQFNGFNCDANFHSRFPAERDISVYCGVQTITLKINFCPVLF 60
Salmon          MEQICLIILLISKALSVGAQFNGYNCDANFHSRFPAERDISVYCGVQTITLKINFCPVLF
Human           MEQIWLLLLLTIRVLPGSAQFNGYNCDANLHSRFPAERDISVYCGVQAITMKINFCTVLF
                **:: :::**  :.:  .*****:*****:****************:**:*****.***

Zebrafish       SGYTDTDLALNGRHGDAHCRGFINNNTFPTVVLFSISLNTLESCGNSLVVSTAHGPNAYG 120
Salmon          SGYTDTDLALNGRHGDAHCRGFINNNTFPTVVLFSISLATLETCGNALVVSTAQGPNAYG
Human           SGYSETDLALNGRHGDSHCRGFINNNTFPAVVIFIINLSTLEGCGNNLVVSTIPGVSAYG
                ***::**********:**********:**:* *.* *** *** *****  * .***

Zebrafish       NLSLVQIGNIAGYIDTPDPPTVISYLPGLLYKFSCSYPLEYLVNNTQLASSAAAISVKDS 180
Salmon          NLSLVQIGNISGYIDTPDPPTIISYLPGLLYKFSCSYPLEYLVNNTQLASSAAAISVKDS
Human           NATSVQVGNISGYIDTPDPPTIISYLPGLLYKFSCSYPLEYLVNNTQLASSSAAISVREN
                * : **:***:*********:**************************:*****::.

Zebrafish       NGTFVSTLNLLLYNDSSYIQNLAIPMSGLTLKTRVFAAVKATNLDRRWNVLMDYCYTTPS 240
Salmon          NGTFVSTLSLLLYNDSSYVQQLSIPMAGLTLKTRVFAAVKATNLDRRWNVLMDYCYTTAS
Human           NGTFVSTLNLLLYNDSTYNQQLIIPSIGLPLKTKVFAAVQATNLDGRWNVLMDYCYTTPS
                ********.*******:* *:* **  **.***:*****:***** ***********.*

Zebrafish       GNPNDELRYDLFFGCDKDPQTTVFENGKSQMGRFSFEVFRFVKHKNQKMSTVFLHCVTKL 300
Salmon          GNPNDELRYDLFFSCDKDPQTTVFENGKSQMGRFAFEVFRFVKHKNQKMSTVFLHCVTKL
Human           GNPNDDIRYDLFLSCDKDPQTTVIENGRSQRGRFSFEVFRFVKHKNQKMSTVFLHCVTKL
                *****::*****:.*********:***:** ***:************************

Zebrafish       CRSDDCPMLLPICGKRKKRDVSERS----GGASDNAVITAGPIITRSDETPSNISQLAQL 356
Salmon          CRADDCPMLMPICGSRKKRDVSERTES--NSASGNAIITAGPIITRSDDTPMNNSQLAQL
Human           CRADDCPFLMPICSHRERRDAGRRTTWSPQSSSGSAVLSAGPIITRSDETPTNNSQLGSP
                **:****:*:***. *::**...*:     .:*.*:::*********:** * ***..

Zebrafish       NGPPFKMNSVTSALISAIAILAVMSMCFFILSMSLLRGKQTPATSLSGSHNPAFS  411
Salmon          NSPVFRMNTVTSALISGIIILGVMSLCFFILSLTLLKGKRAPPTILSGARNPAFN
Human           SMPPFQLNAITSALISGMVILGVTSFSLLLCSLALLHRKGPTSLVLNGIRNPVFD
                . * *::*::******.: **.* *:.::: *::**: * ...  *.* :**.*.
```

Figure 2. Visual multiple sequence alignment using CLUSTAL W (https://www.genome.jp/tools-bin/clustalw) (Thompson et al., 1994). Zebrafish (accession A0A0R4ITH5), salmon (accession ACN10635, ~86.4% identity) and human (accession AAH31261, ~71.5% identity) ZPLD1 amino acid sequences are highly conserved.

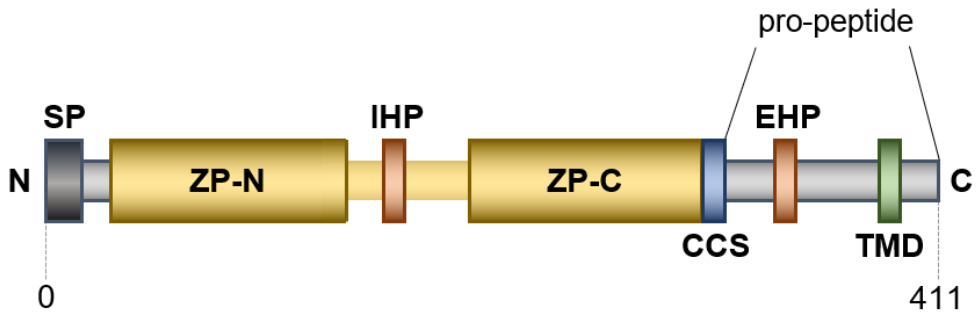


Figure 3. Scheme of zebrafish ZPLD1 structure expected to contain a N-terminal signal peptide, a bipartite Zona Pellucida (ZP) module (including ZP-N and ZP-C domains) and a C-terminal pro-peptide (Dernedde et al., 2014). The internal (IHP) and external (EHP) hydrophobic patches, the consensus cleavage site (CCS) and the transmembrane domain (TMD) are also indicated.

Accordingly, I aim to investigate if ZPLD1, a protein which does not display a high identity to classical ZP proteins, is able to multimerize and therefore likely accounts for polymeric assembly of the gelatinous cupula. Furthermore, I want to assess if we can identify the regulatory hydrophobic peptide regions based on their conserved structure and amino acid characteristics to understand their role for ZPLD1 maturation. I want to confirm that cleavage at the CCS as well as membrane anchoring are indeed prerequisites for proper polymerization and gather first hints on the protease required for ZPLD1 maturation and corresponding activation.

2 Materials and methods

2.1 Protein sequence analyses

Sequence homologies as well as visual multiple sequence alignments were calculated and created using CLUSTALW (https://www.genome.jp/tools-bin/clustalw) (Thompson et al., 1994). Secondary structure prediction was carried out using PSIPred (http://bioinf.cs.ucl.ac.uk/psipred/) (McGuffin et al., 2000), Jpred4 (https://www.compbio.dundee.ac.uk/jpred/) (Drozdetskiy et al., 2015), Porter 5.0 (http://distilldeep.ucd.ie/porter/) (Pollastri and McLysaght, 2005), Sspro (http://download.igb.uci.edu/sspro4.html) (Cheng et al., 2005), and PredictProtein (https://predictprotein.org/) (Rost and Liu, 2003). Position of the TMD was predicted using TMHMM 2.0 software (Krogh et al., 2001). Position of the CFCS was predicted using the SignalP algorithm (Petersen et al., 2011). Structure of zebrafish ZPLD1 was predicted by AlphaFold (AlphaFold DB version 2022-06-01) and visualized using pyMOL (The PyMOL Molecular Graphics System, Version 2.5.4 Schrödinger, LLC.) ("AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models | Nucleic Acids Research | Oxford Academic," n.d.; "PyMOL | pymol.org," n.d.; Jumper et al., 2021).

2.2 ZPLD1 expression constructs

Wild-type zebrafish ZPLD1 DNA cloned into vector pEGFP-N1 was purchased from General Biosystems. The full sequence can be found in the supplementary information (Supplementary Figure 1). All genetic manipulations (deletion and insertions) were performed using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) in combination with a back-to-back primer design. The back-to-back primer design allowed for deletions generated simply by positioning both 5' ends of forward and reverse primers directly on the sequence flanking the desired deletion. For insertions, additional DNA was added to the 5' ends of forward and/or reverse primers. A linear double strand was generated while the vector was used as a template. In a subsequent reaction

phosphorylation at 5´ends and vector ligation was performed. Addition of enzyme DpnI cut methylated template DNA and therefore enhanced transformation efficiency. Primers were ordered from Metabion international AG. Primer sequences can be found in Table 1. The Strep- tag was inserted between amino acids K_{315} and R_{316} with a short SSGS linker anterior to the Strep sequence. The putative EHP was deleted between amino acids V_{333} and R_{342}. The putative IHP was deleted between amino acids E_{360} and T_{366}. The TMD domain was deleted between amino acids V_{366} and S_{390}. The CCS was deleted between amino acids R_{316} and R_{319}. The 8xHis-tag was inserted between amino acids Q_{20} and F_{21} with a short SSGS linker posterior to the 8xHis-tag (N-His), between amino acids S_{325} and G_{326} (C-His1), and in-between S_{348} and N_{349} (C-His2) respectively. All constructs were sequence-verified before transfection.

Table 1: Primers (5´-3´) used to generate ZPLD1 expression constructs

Mutation	**Primer**
Deletion of eGFP	Forward: CAGCCATACCACATTTGTAGAGGTTTTACTTGC
	Reverse: TCACTCGAGGCTGAACGCTGGG
Insertion of the Strepavidin-tag (and corresponding Linker)	Forward: TGGAGCCATCCGCAGTTTGAAAAAAGGAAGAAAAGGGATGTGTCA GAAAGAAGCGG
	Reverse: AGAGCCAGAAGACTTACCGCAGATGGGCAGGAGCATGGGGC
Deletion of the putative EHP	Forward: AGTGATGAAACTCCCAGTAATATATCCC
	Reverse: TGCATTGTCTGAGGCACCACCGC
Deletion of the putative IHP	Forward: CAGCTTGCTTCGTCGGCCGCAGCG
	Reverse: CAGTGGGTAGCTACAACTGAACTTGTAAAGCAGTCC
Deletion of the putative TMD	Forward: ATGTCGCTGCTGAGGGGAAAACAGACG
	Reverse: TGAATTCATTTTGAATGGAGGGCCG
Deletion of the putative CCS	Forward: CCATCTGCGGTAAGGCTAAGAAAGCTGATGTGTCAGAAGGAAGCGG
	Reverse: GCAGGAGCATGGGGCAGTCATCAGAGCG
Insertion of the N-terminal 8xHis-tag (N-His)	Forward: TCTTCTGGCTCTTTCAATGGGTTCAACTGTGATGCCAACTTCC
	Reverse: ATGATGATGATGGTGATGGTGATGTTGAGCATTCGCTATGAAAGTTT TACTTACAAGC
Insertion of the C-terminal 8xHis-tag N-terminal of the EHP (C-His1)	Forward: CATCATCATCATGGTGGTGCCTCAGAC
	Reverse: GTGATGGTGATGGCTTCTTTCTGACACATCCC
Insertion of the C-terminal 8xHis-tag C-terminal of the EHP (C-His2)	Forward: CATCATCATCATAATATATCCCAACTAGCGCAG
	Reverse: GTGATGGTGATGACTGGGAGTTTCATCACTTCG

2.3 Cell lines and culture conditions

Mardin-Darby canine kidney (MDCK NBL-2) cells were purchased from ATCC and grown in DMEM supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 U/ml streptomycin, 25 mM D-glucose at 37 °C, 5% CO_2. Stable clones were generated by transfecting MDCK cells using Lipofectamine 3000 (Invitrogen) following the manufacturer`s protocol. Selection was started 24 h after transfection by adding 1 mg/ml Geniticin (G418 sulfate) (Invitrogen) and was pursued for 1-3 weeks to obtain a G418 resistant cell population.

2.4 Protein purification and Western blot

MDCK cells stably expressing Strep-tagged ZPLD1 were grown on 100 mm polystyrene cell culture dishes (Corning). As soon as cells reached confluence the medium was collected and stored at -20 °C until further use for purification.

1 ml of Strep-tactin-slurry (Iba Lifesciences) was transferred to a reusable gravity flow column. Collected medium was thawed, centrifuged for 20 min at 1400 x g to spin down any contained cells and cell components and pH measured using a pH test strip (Merck). Pooled media with a pH between 7 and 8 was run over the column and purification conducted according to the manufacturer´s protocol.

Elution fractions were pooled, concentrated using an Amicon Ultra 0.5 ml 10 K centrifugal filter device (Merck), measured at A280 (NanoDrop) and stored at 4 °C.

Purified ZPLD1 was deglycosylated under denaturing conditions using PNGase F (New England Biolabs) according to the manufacturer´s protocol. Deglycosylated protein was reduced in Laemmli buffer containing 2.5% or 5% mercaptoethanol at 96 °C for 5 or 15 min and applied to a 12% SDS-polyacrylamide gel electrophoresis (PAGE). After blotting the nitrocellulose membrane (GE Healthcare) was incubated with polyclonal guinea pig antibodies (1:10,000 dilution) against two distinct ZPLD1 peptide sequences, as previously described (Dernedde et al., 2014). A secondary horseradish peroxidase-conjugated antibody (1:4,000 dilution, Jackson ImmunoResearch) was subsequently applied for visualization.

A Western blot detecting untreated ZPLD1 as well as a negative control mock transfected sample can be found in Supplementary Figure 2.

2.5 Treatment of cells with protease inhibitors

Stably transfected MDCK cells were transferred to OptiMEM (without supplements) split to, and subsequently grown on 8 well Lab-Tek II chamber slides (Thermo Scientific). For protease inhibitor cocktail (PIC) treatment cells were grown for 6 h before addition of 0.1%, 0.15%, 0.2% of PIC or 0.2% DMSO respectively. For single proteases treatments 0.1 mM AEBSF, 0.2 mM AEBSF, 1.4 µM E64, 4 µM bestatin, 1.5 µM pepstatin A, and 0.2 % DMSO were added right after splitting cells to chamber slide wells. All media including supplements were renewed every ~ 24 hr. Immunofluorescence analysis was performed after ~ 48 h.

2.6 Immunofluorescence

Cells grown on 8 well Lab-Tek II chamber slides (Thermo Scientific) were fixed in 4% paraformaldehyde (PFA) for 30 min. When needed, cells were permeabilized 30 min at room temperature with 0.5% Triton X-100. After washing in PBS, cells were incubated 30 min at room temperature in 10% goat serum in PBS. Cells were then incubated for 2 h at room temperature with the mouse anti-strep-tag II monoclonal antibody (1:100 dilution; Iba Lifesciences) and where needed simultaneously with the rabbit anti-His-tag antibody (1:400 dilution; Invitrogen). If needed, permeabilized cells were simultaneously co-stained with rabbit anti-GM130 Golgi family protein antibody (1:100 dilution; Abcam). Cells were washed in PBS and incubated for 1 h at room temperature with the appropriate secondary antibody: Alexa-Fluor 594-conjugated goat secondary antibody against rabbit immunoglobulin G (IgG) (dilution 1:400; Abcam); or Alexa-Fluor 488-conjugated goat antibody against mouse IgG (1:200 dilution; Abcam). Cells were then stained for 5 min with 4,6-diamidino-2-phenylindole (DAPI) and mounted using Shandon ImmuMount (Thermo Scientific). All slides were visualized under a Zeiss

AxioCam MRm fluorescence microscope. Images were evaluated with the AxioVision 4.8.2 SP3 program.

Quantification of ZPLD1 polymer area was performed based on nine independent pictures for each condition taken at 10x magnification. Pictures were adjusted for brightness and contrast and analyzed using the ImageJ software (https://imagej.net/ij/index.html) ("ImageJ," n.d.; Schneider et al., 2012). After setting a threshold to exclude the membrane signal background, polymers were quantified with an object size limit of 300. Data are shown as average ± SEM and significant relevance measured by unpaired two-tailed *t*-test analysis (https://www.mathportal.org/calculators/statistics-calculator/t-test-calculator.php) ("T-Test Calculator with step by step explanation," n.d.).

Control immunofluorescence images of untransfected MDCK cells can be found in the supplementary materials (Supplementary Figure 3).

3 Results

3.1 Transformed MDCK cells express physiological active ZPLD1

To confirm that ZPLD1 is able to form homopolymers and to test any further hypotheses, a suitable cellular expression system is wanted. Finally, I transfected Mardin-Darby Canine Kidney (MDCK) cells that have been previously shown to express, process and secrete the polymer forming ZP protein Uromodulin in an active form (Schaeffer et al., 2009). To facilitate ZPLD1 protein purification a Strep-tag with a short linker sequence was inserted 5´ to the conserved CCS between amino acids K_{315} and R_{316}. Transformed MDCK cells secreted Strep-tagged ZPLD1. The protein was further purified by affinity chromatography from the culture supernatant and separated by gel-electrophoresis after enzymatic removal of N-glycans. ZPLD1 migrates at about 35 kDa similar to ZPLD1 extracted from salmon cupulae (Figure 4 A) (Dernedde et al., 2014). Obviously, the tag insert has no adverse effect on protein maturation. ZPLD1 assembles into extracellular polymers on the surface of transfected MDCK cells as observed by immunofluorescence analyses (Figure 4 B). Visible polymeric structures appear as bundles or matrices likely composed of single filaments.

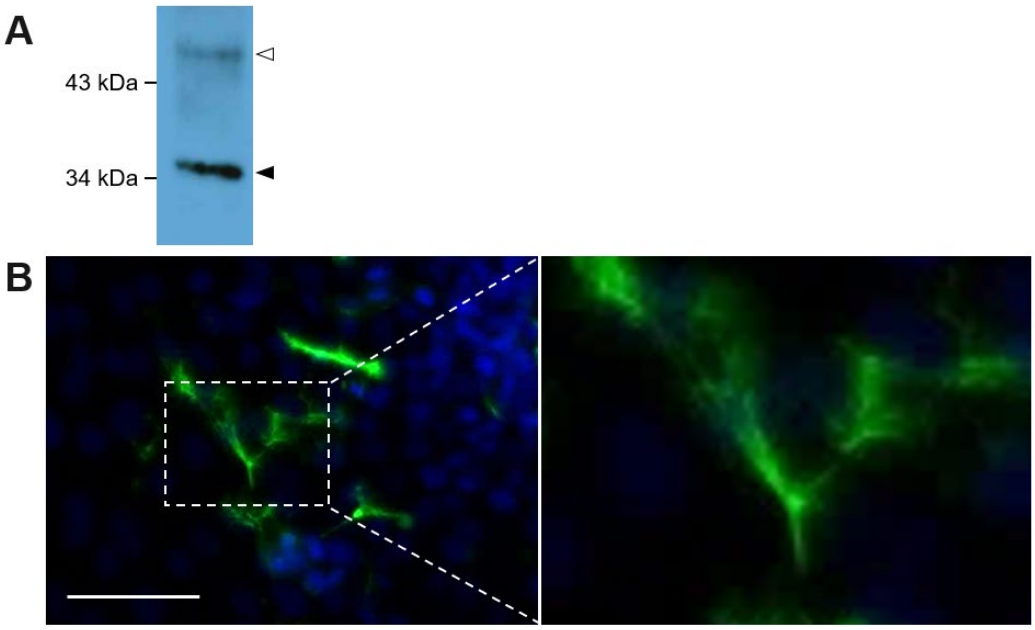


Figure 4. MDCK cells as a model to study ZPLD1. (A) Western blot of PNGase F treated ZPLD1, purified from cell culture media of transfected MDCK cells. ZPLD1 still containing N-glycans (white arrowhead) detected

at higher molecular weight as a fuzzy protein band compared to the discrete fully deglycosylated protein (black arrowhead); (B) Immunofluorescence image of non-permeabilized MDCK cells expressing Strep-tagged ZPLD1 stained with anti-Strep antibodies. Polymers formed by the protein are clearly detected on the cell surface, likely consisting of matrices or bundles of single filaments; scale bar, 50 µm.

3.2 ZPLD1 proteins lacking the EHP or IHP are not properly processed

The EHP and IHP motifs are reported to be conserved in different ZP proteins regarding their secondary structure but also, to a lesser extent, with respect to the consecutive composition of hydrophobic, aliphatic, small, turnlike and polar amino acids (Schaeffer et al., 2009). Based on these characteristics the amino acid sequence of ZPLD1 was analyzed. For the secondary structure prediction five different computational programs were applied (Figure 5). There is exactly one sequence area ($_{333}$VITAGPIITR$_{342}$) at the C-terminus of ZPLD1 that matches the corresponding consensus grouping set sequence (..sGPl…) for the EHP. In line with the reported high degree of secondary structure conservation in ZP proteins, the putative EHP of ZPLD1 configures two conserved β-strands within an unstructured region between the cleavage site (CCS) and the membrane anchoring domain (TMD) (Schaeffer et al., 2009). For the IHP two overlapping sequence areas ($_{260}$EYLVNNT$_{266}$ and $_{265}$NTQLASS$_{271}$) were identified, that match the consensus grouping set sequence (p..ls.t). However, only the N-terminal sequence seems to form a hydrophobic β-strand as suggested feature of secondary structure conservation (Schaeffer et al., 2009). I therefore propose here the allocation of the IHP of ZPLD1.

Mutagenesis studies were performed to prove whether the identified sequence regions indeed correspond to the hydrophobic motifs and to better understand if their function for protein polymerization resembles the mechanism described for other ZP proteins.

Deletion constructs (Figure 6 A, B) lacking the entire respective sequences were expressed in MDCK cells and phenotypically characterized via microscopy (Figure 6 C). ZPLD1 deficient of its putative EHP does not form extracellular filaments and exhibit impaired secretion. The protein seems to stick in the cytoplasmic membrane without polymerizing (Figure 6 C). For the variant lacking the putative IHP motif, polymer

formation is impaired as well. Only abnormally short and less structured filaments might be suspected at the plasma membrane of the cells. Secretion of the variant lacking the putative IHP sequence seems not to be impaired as compared to the variant lacking the EHP. These findings support our hypotheses regarding the positioning, sequence, and function of ZPLD1´s hydrophobic patches and are in line with the processing of Uromodulin (Schaeffer et al., 2009).

3.3 Depletion of EHP or IHP leads to the formation of intracellular polymeric structures

Schaeffer et al. described intracellular polymer formation for soluble Uromodulin lacking one of its hydrophobic patches and its membrane anchor (Schaeffer et al., 2009). This finding supported the hypothesis that the EHP and IHP play an important role for regulation and prevention of intracellular polymerization thereby enabling proper processing and secretion. To proof this for ZPLD1 we created an EHP mutant lacking the TMD domain (Figure 6 A) and observed the same phenotype with intracellular polymeric structures (Figure 7 A).

Here it is interesting to note that also the isoforms still containing the TMD seem to produce the same kind of intracellular filament structures. Schaeffer et al. could not show colocalization of the intracellular filaments of Uromodulin with the ER marker Calnexin (Schaeffer et al., 2009). For ZPLD1 mutants we further checked if an overlapping staining of ZPLD1 with an anti-GM130 antibody, a Golgi marker, is detectable. This was not the case, and we therefore conclude that the observed intracellular ZPLD1 complexes are no longer part of the vesicular transport system (Figure 7 B).

3.4 Release of the C-terminal pro-peptide and hence loss of the putative EHP is a prerequisite for proper extracellular polymerization

As already suggested for other ZP proteins and here for ZPLD1 it seems likely that the two hydrophobic motifs are functionally related and necessary to keep the ZP module in an inactive conformation that prevents intracellular polymerization. Next, I assessed whether this regulatory function is lost upon cleavage at the CCS that releases the EHP motif. To address this question, I generated different insertions to the wild-type ZPLD1 sequence generating an N-terminal His-tag (N-His), a C-terminal His-tag upstream of the EHP motif (between S_{325} and G_{326}) (C-His1), and another variant carrying a C-terminal His-tag downstream of the EHP motif (between S_{348} and N_{349}) (C-His2) (Figure 8). These isoforms were as well stably expressed in MDCK cells.

Immunofluorescence experiments with nonpermeabilized cells showed the presence of polymers at the plasma membrane of all constructs when using an antibody against the Strep-tag positioned immediately upstream of the CCS. These polymers looked identical to the ones observed in stable clones expressing untagged wild type ZPLD1, indicating that the presence of N- and C-terminal His-tags do not interfere with protein processing and assembly into filaments. Only polymers containing the N-His isoform were positive for an anti-His antibody, suggesting that sequence downstream of S_{325}, i.e., including the EHP motif, is lost in polymeric ZPLD1.

These findings support the hypothesis, that ZPLD1´s EHP and IHP motifs have the same function as described for other ZP proteins and that loss of the EHP motif is essential for ZPLD1 as for other ZP-proteins to enable polymerization into extracellular filaments.

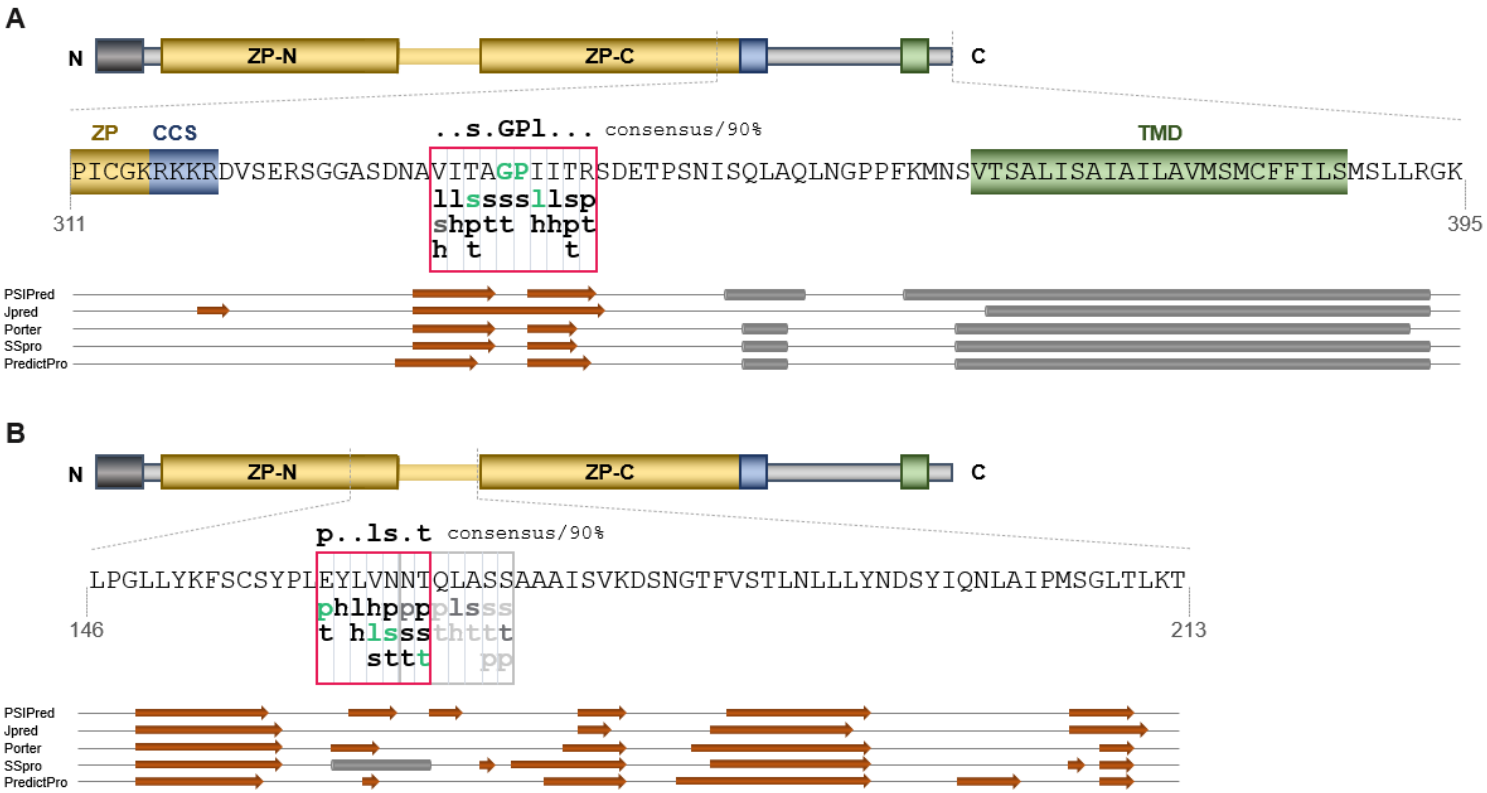


Figure 5. Identification of the putative IHP and EHP motifs within in ZPLD1 sequence. scheme, detailed amino acid sequences, and annotated secondary structure predictions are shown for the EHP (A) and IHP (B). The signal peptide is shown as a dark grey box, the ZP module is shown in yellow and is likely divided in ZP-N and ZP-C domains. The consensus protease cleavage site (CCS) is shown in blue and the transmembrane domain (TMD) in green. The predicted hydrophobic patches are framed in red. Consensus with the amino acid grouping-set sequences for EHP (. . s . GPl . . . / consensus 90%) and IHP (p . . ls . t / consensus 90%) described by Schaeffer et al. is indicated as green lowercase letters (Schaeffer, 2009). The lowercase letters indicate the following grouping-sets: l, aliphatic (I, L, V); s, small (A, C, D, G, N, P, S, T, V); h, hydrophobic (F, H, I, L, M, V, W, Y); p, polar (C, D, E, H, K, N, Q, R, S, T); and t, turn like (A, C, D, E, G, H, K, N, Q, R, S, T). Secondary structure predictions were performed with five different programs (for details see materials and methods section). β-strands are depicted as orange arrows, α-helices as gray cylinders.

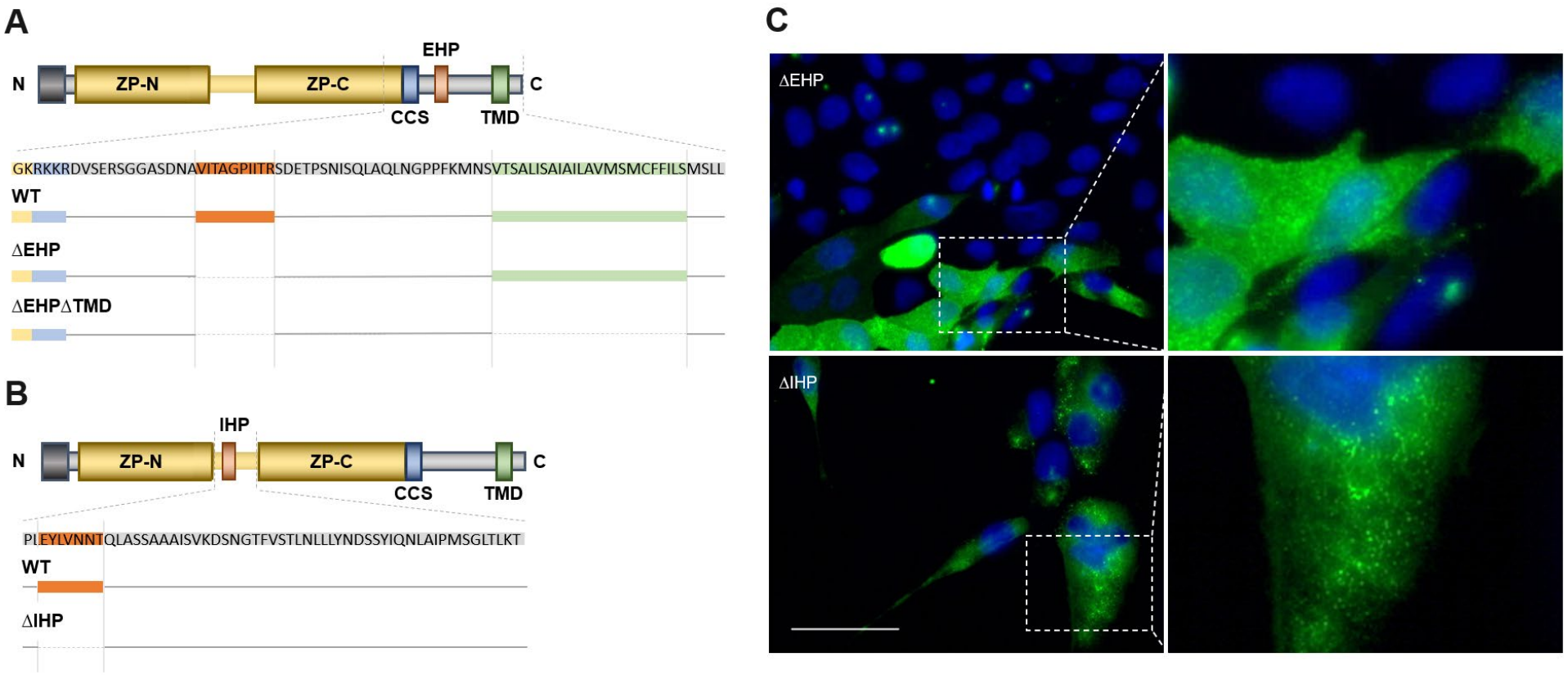


Figure 6. Polymerization defect in mutants lacking putative EHP and IHP sequences. (A + B) domain structure and sequence excerpt of ZPLD1. The putative EHP and IHP motifs are highlighted in red, and they were completely deleted in the ΔEHP and ΔIHP variants. WT: wild type; (C) Immunofluorescence analysis showing mutant ZPLD1 on the cell surface of stably transfected MDCK cells fixed with PFA. ZPLD1 from both variants is capable to traffic to the plasma membrane, but unlike the wild type protein it does not polymerize properly. ΔIHP forms abnormally short and less organized polymers while ΔEHP does not form polymeric structures at all and seems to be stuck unprocessed in the plasma membrane. Scale bar, 50 µm.

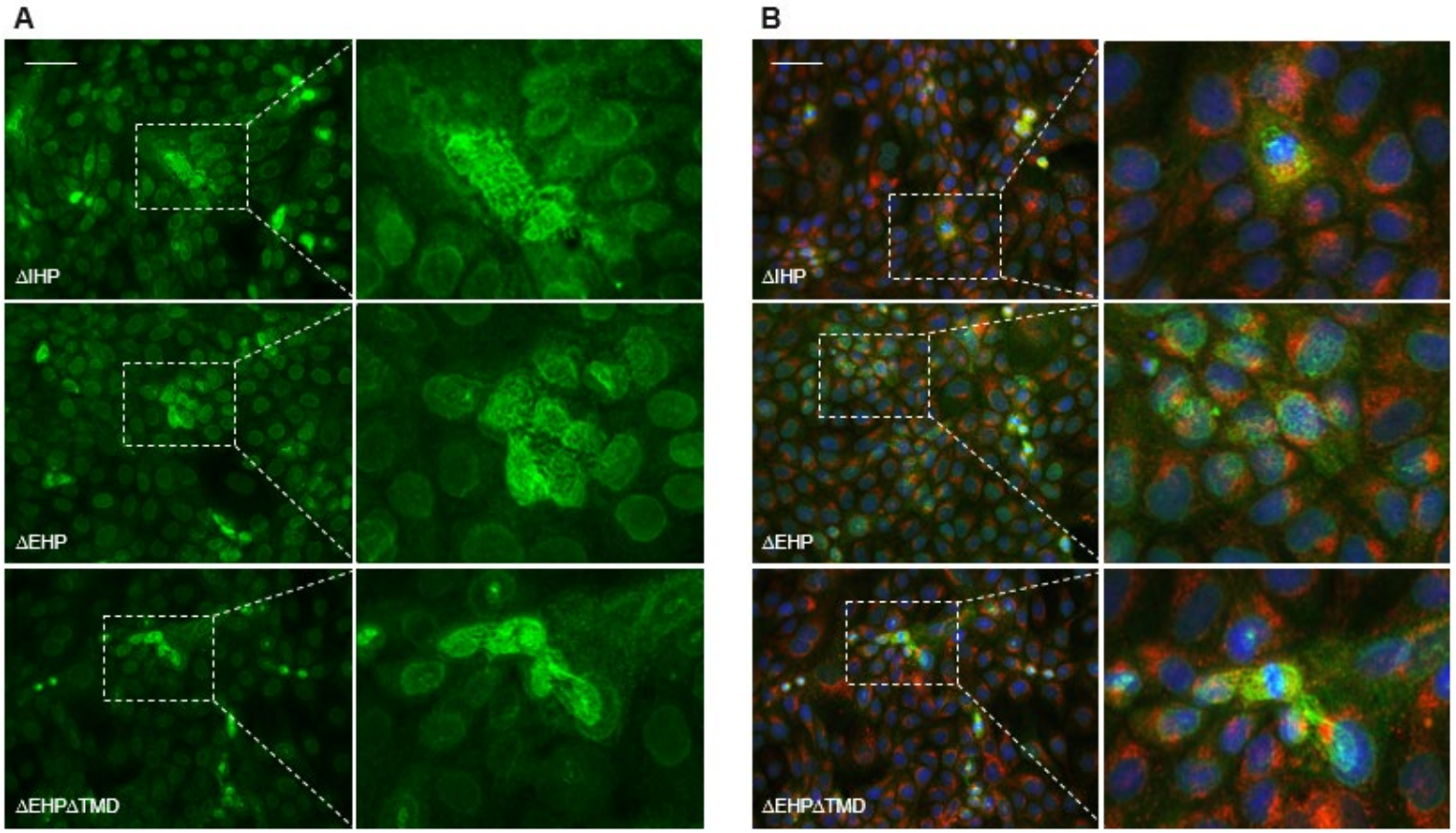


Figure 7. Deletion of the EHP or IHP lead to the formation of intra-cellular polymers. (A) Immunofluorescence analysis of permeabilized MDCK cells stably expressing different ZPLD1 mutant isoforms. In contrast to findings for Uromodulin (Schaeffer et al., 2009), intracellular assembly of mutant protein can be observed for all three investigated variants; Bar, 50 µm; (B) Immunofluorescence of MDCK cells stably expressing ZPLD1 IHP and EHP mutant proteins. Permeabilized cells were incubated with anti-Strep, anti-GM130 Golgi marker and DAPI. The presented merged pictures show Strep (ZPLD1) and GM130 signals in green and red, respectively. DAPI cell nucleus staining is shown in blue. Intracellular filaments formed by the mutant protein do not colocalize with GM130; Bar, 50 µm.

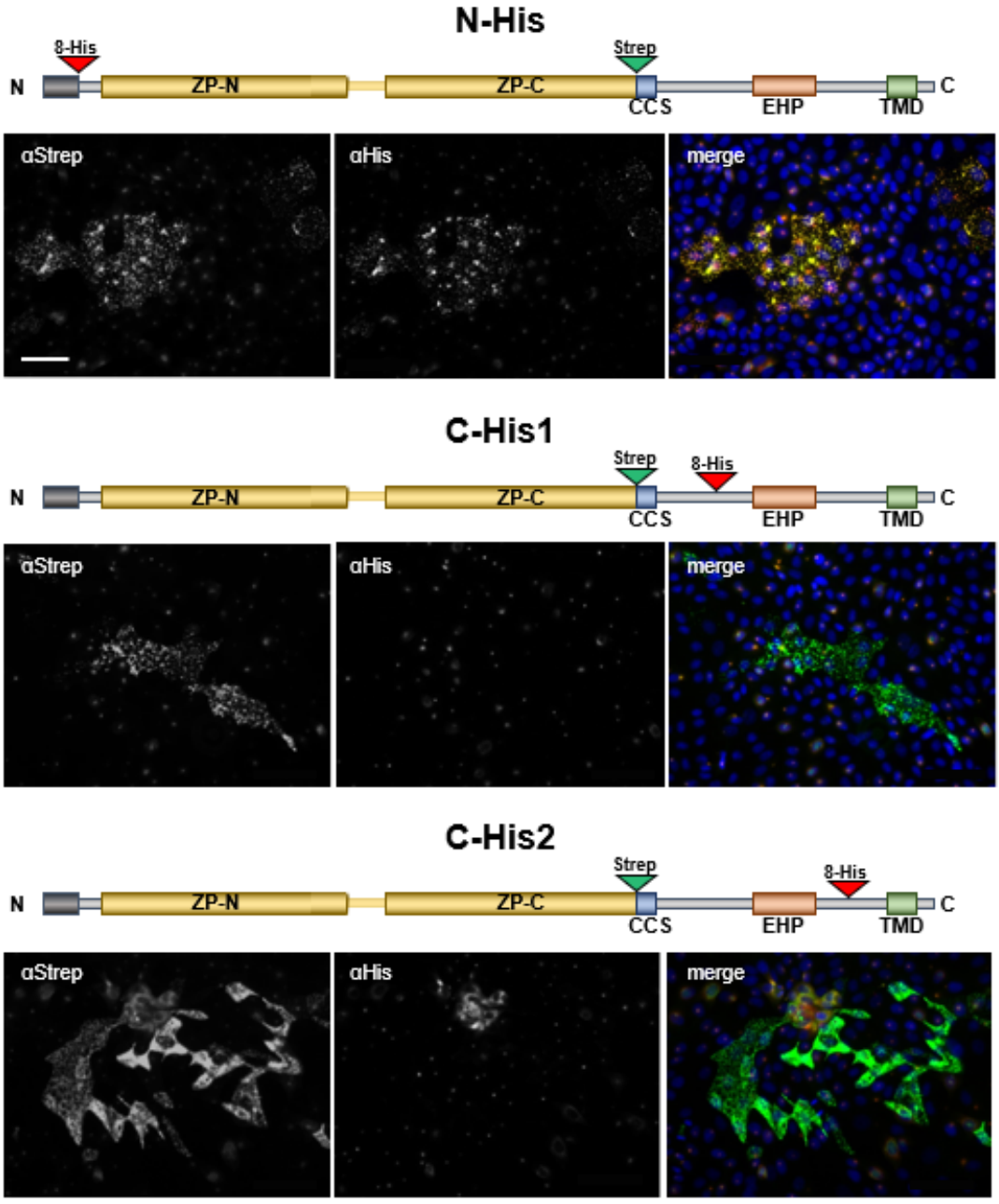


Figure 8. The EHP motif is lost upon ZPLD1 maturation and assembly into filaments. Immunofluorescence analysis on MDCK cells stably expressing N- and C-terminally tagged ZPLD1. A schematic representation of each tagged isoform is shown above the respective immunofluorescence panel. ZPLD1 structure is depicted as before. The EHP motif is indicated as an orange box, the position of 8xHis and Strep-tags as red and green triangles, respectively. Immunofluorescence analysis was carried out on unpermeabilized PFA-fixed cells that were stained for the His-tag and the Strep- tag simultaneously. In the merged picture, the Strep-tag signal is shown in green whereas the 8xHis-tag signal is shown in red. ZPLD1 polymers are positive for both antibodies when the 8xHis-tag is present at the N terminus (N-His). However, the His epitope is not detectable when located downstream of the CCS, either before (C-His1) or after (C-His2) the EHP motif. These data suggest that the EHP sequence is lost in polymeric ZPLD1 proteins. Bar, 50 µm.

3.5 Identified CCS as well as membrane anchoring is required for processing at the plasma membrane

To enable above-described loss of the EHP, a maturation process including proteolytic cleavage of the pro-peptide is required. As mentioned earlier, I hypothesized that proteolytic cleavage happens at the putative CCS at position $_{315}$RKKR$_{318}$ and requires membrane anchoring by a TMD. The sequence and positioning of the transmembrane domain $_{385}$VTSALISAIAILAVMSMCFFILS$_{407}$ was predicted using a computational software program. To test these hypotheses two further ZPLD1 expression constructs, encoding for a variant lacking the putative CCS (ΔCCS) as well as a variant lacking its TMD (ΔTMD) were designed (Figure 9 A).

As expected, when expressed in MDCK cells, both ZPLD1 variants lack their ability to form polymers. ZPLD1 deleted for its CCS seems to be locked in the plasma membrane (Figure 9 B). Exactely the same picture has been reported for Uromodulin, carrying a mutation of its cleavage site (Brunati et al., 2015). This observation suggests that separation of its pro-peptide containing the TMD is a prerequisite for ultimate secretion of mature ZPLD1. ZPLD1 being exempt from its TMD cannot be detected on nonpermeabilized cells (Figure 9 B). Proper transfection and expression can be proven however, by analyzing permeabilized cells transfected with the ΔTMD construct (Figure 9 C). For this variant as well, experiments on the ZP protein Uromodulin resulted in exactly the same outcomes (Brunati et al., 2015).

These results indicate that proteolytic cleavage is only required for secretion as membrane anchoring by the transmembrane domain needs to be overcome. Without the TMD, ZPLD1 is secreted from the cells but does not polymerize. Furthermore, this observation clearly supports the hypothesis that membrane anchoring is required for ZPLD1 polymerization. However, it cannot give a clear hint if this requirement is explained by e.g., the type and location of the protease responsible for cleavage, steric effects and required proximity of single ZPLD1 molecules during polymerization process, or any other not yet considered reason.

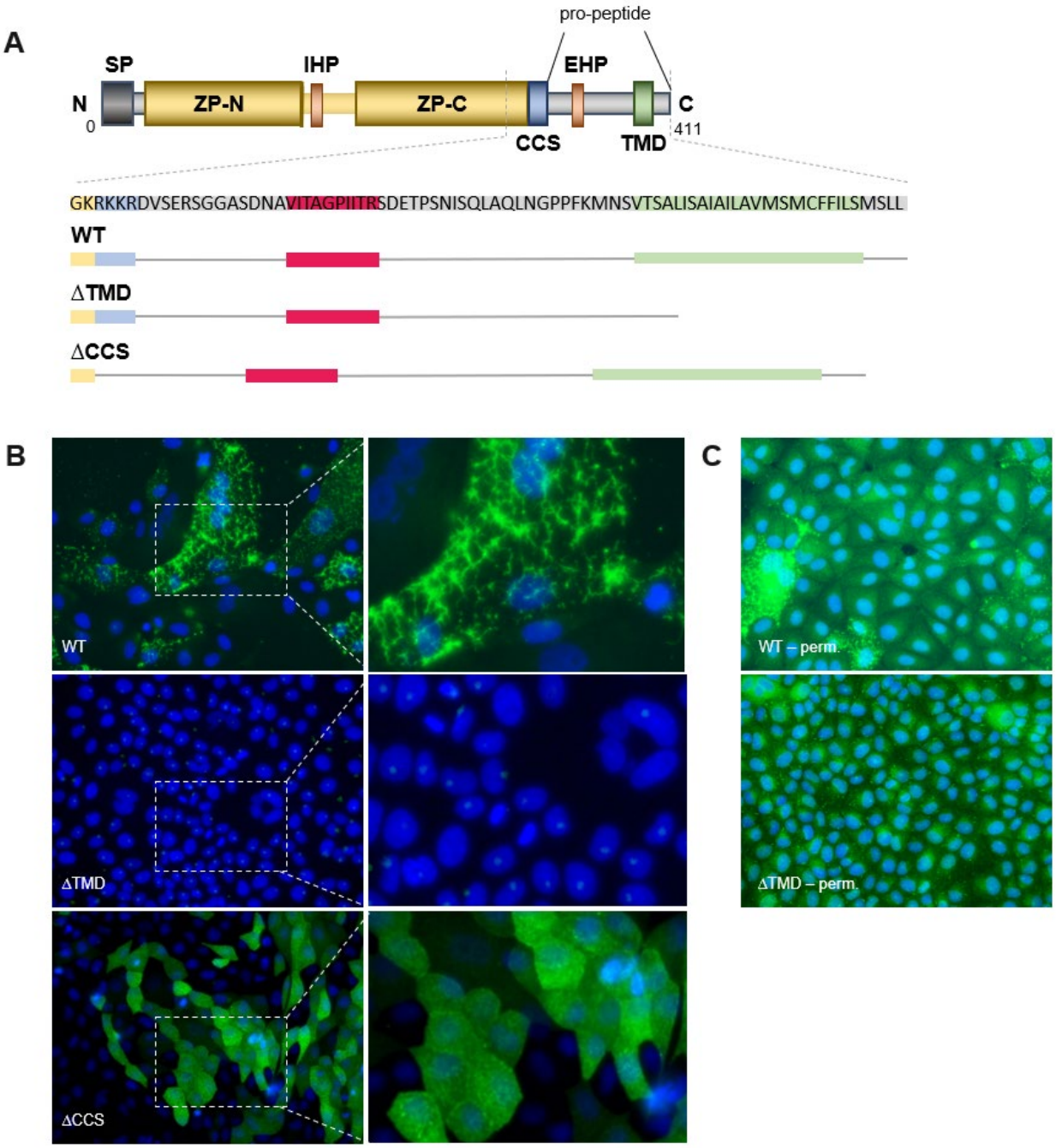


Figure 9. Polymerization defect in mutants lacking putative TMD and CCS sequences. (A) domain structure and sequence excerpt of ZPLD1. The putative TMD and CCS motifs are highlighted in green and blue respectively and were completely deleted in the ΔTMD and ΔIHP variants. WT: wild type; (B) Immunofluorescence analysis showing mutant ZPLD1 on the cell surface of stably transfected MDCK cells fixed with PFA. Both variants are expressed, but unlike wild type protein they do not polymerize properly. ΔTMD gets secreted and cannot be detected at the plasma membrane, while ΔCCS seems to be locked unprocessed in the plasma membrane. (C) Permeabilized MDCK cells transfected with the ΔTMD and WT ZPLD1 variant, are both expressed to a similar extend and visible inside of the cells. Scale bar, 50 µm

3.6 Cleavage of ZPLD1 is dependent on a serine type protease

To gain more insights on the type of protease responsible for the proteolytic maturation process, it was initially tested if polymerization can be altered by the addition of protease inhibitors to transfected MDCK cells. The treatment of ZPLD1-expressing MDCK cells with a protease inhibitor cocktail (PIC) indeed led to significant reduction of observed polymerization on the membrane surface (Figure 10 A). Cells of the exact same passage and sample were treated with an equivalent amount of DMSO as suitable control. Thereby, alteration of protein expression or intracellular distribution was excluded. The clear dose dependency of polymerization inhibition further confirms influence of the applied protease inhibitors (Figure 10 B).

When treated with single PIC components, each targeting a specific catalytic type of protease, only AEBSF (4-(2-Aminoethyl) benzene sulfonyl fluoride hydrochloride) significantly reduced ZPLD1 polymerization (Figure 11 A). Doubling the applied amount of AEBSF further reduced the amount of detected polymers, mirroring the clear dose dependency observed in the PIC treatment experiments (Figure 11 B). AEBSF is an irreversible serine protease inhibitor, thereby indicating the requirement of such a serine type protease for the maturation of ZPLD1. Similar results of protease inhibition experiments have been observed for Uromodulin where ultimately Hepsin was identified as processing protease (Brunati et al., 2015).

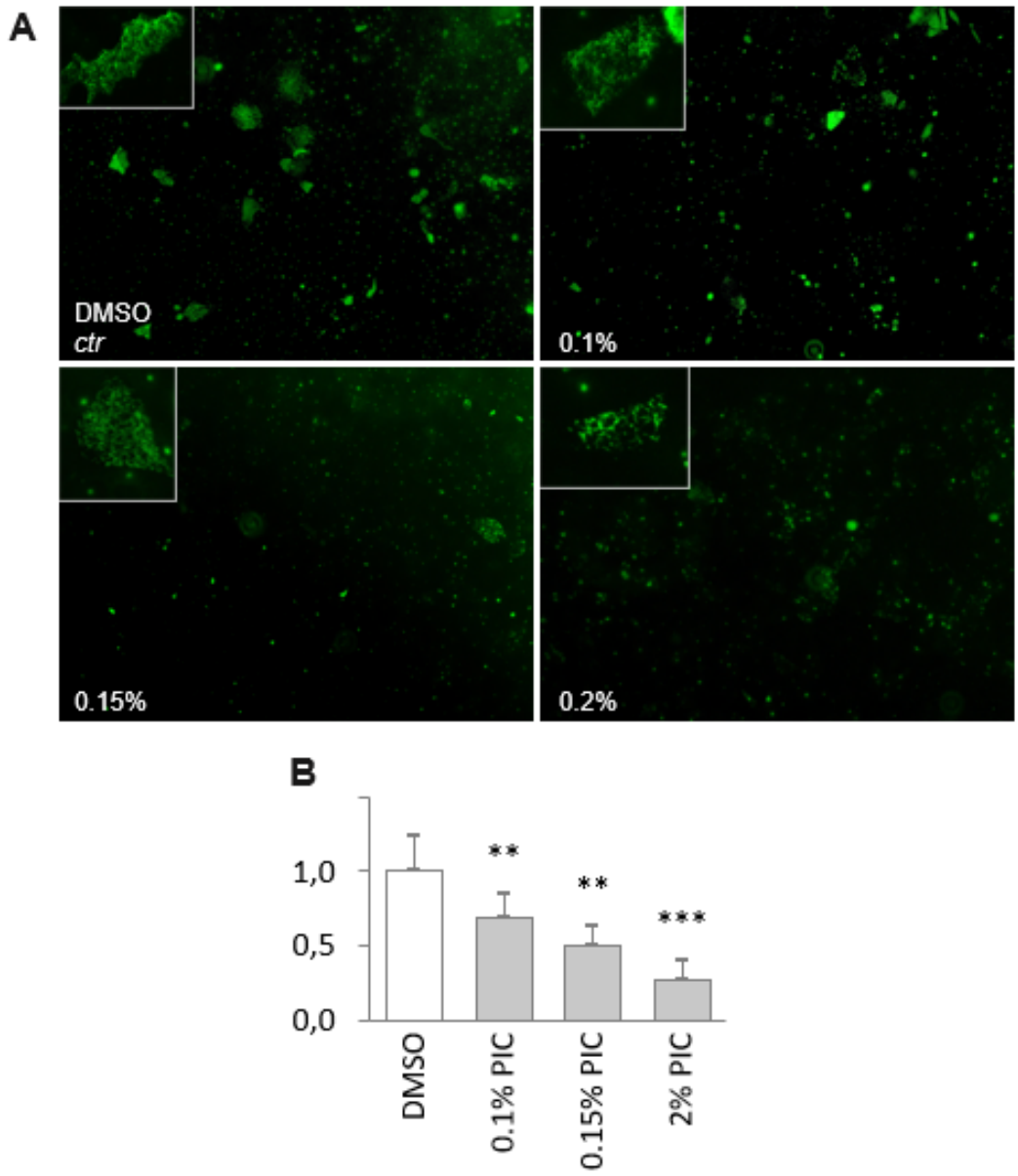


Figure 10. Proteolytic cleavage is required for the release of polymerization competent ZPLD1. (A) Immunofluorescence analysis showing ZPLD1 on the surface of MDCK cells treated with DMSO (control), and different concentrations of protease inhibitor cocktail (PIC). Scale bar, 50 µm. (B) Quantification of the average surface of ZPLD1 polymers shows dose dependent, significant reduction of polymerization on the cell surface. Bars indicate average ± s.e.m. **p<0.01; ***p<0.001 (Mann-Whitney test). The graph represents mean ratios of nine independent immunofluorescence pictures.

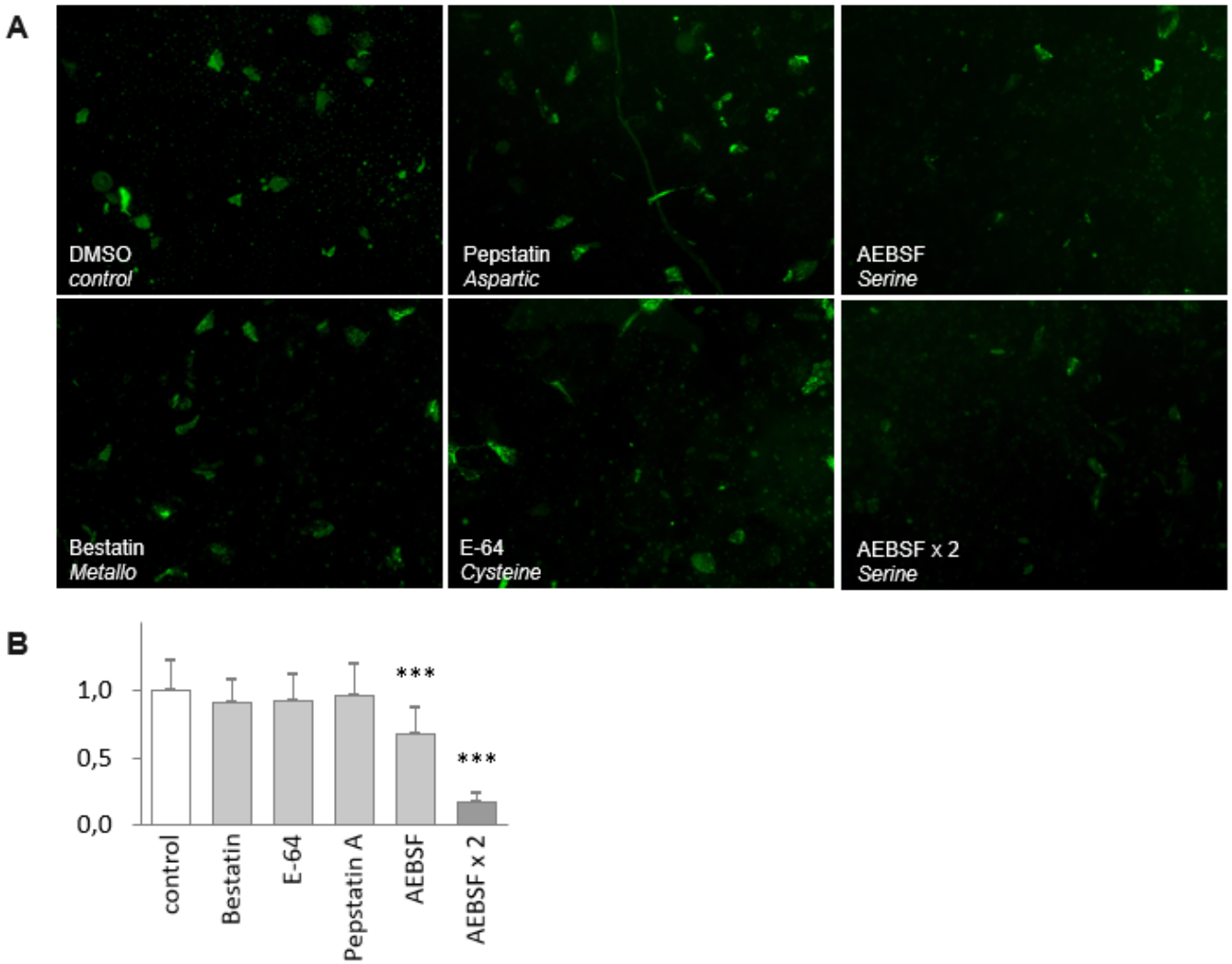


Figure 11. A serine type protease is responsible for the processing of ZPLD1. (A) Immunofluorescence analysis showing ZPLD1 on the surface of MDCK cells treated with DMSO (control), and single protease inhibitors as indicated. Scale bar, 50 µm.

(B) Quantification of the average surface of ZPLD1 polymers shows that only treatment with AEBSF results in dose dependent, significant reduced polymerization on the cell surface. Bars indicate average ± s.e.m. ***p<0.001 (Mann-Whitney test). The graph represents mean ratios of nine independent immunofluorescence pictures.

4 Discussion

4.1 ZPLD1 indeed polymerizes – ZPLD1 as a further homopolymer forming ZP family member

ZPLD1 was identified as a main component of the cupula in the inner ear and right away associated with likely being part of the family of ZP proteins (Dernedde et al., 2014). Most of the members share the ability to polymerize extracellularly into filaments and matrices via their ZP module, while distinct additional functionalities can be attributed (Jovine et al., 2004). In this work, I aimed at investigating if ZPLD1 is able to form homopolymers and gain new insights about its key features as well as the molecular mechanism that regulates this protein processing. Previous studies performed on murine ZP3 protein and recombinant Uromodulin partly served as a blueprint for the experiments regarding ZPLD1 stepwise maturation and polymerization (Brunati et al., 2015; Jovine et al., 2004; Schaeffer et al., 2009; Zhao et al., 2003). A combination of genetic engineering, protein expression in a suitable cell line and immunofluorescence microscopy was necessary to establish a model to assess ZPLD1´s ability to form polymers and gain first insight into the maturation process. The presented work proofs for the first time the formation of ZPLD1 homopolymers together with further characteristics and behaviors described for other ZP-proteins like Uromodulin, α-tectorin (TECTA) or ZP3. Consequently, ZPLD1 indeed seems to be a member of the ZP protein family, likely sharing key conserved structural as well as functional features. Furthermore, this fact greatly supports the hypothesis that ZPLD1 is a major functional component of the cupula hydrogel with great importance for proper vestibular function.

4.2 A matching team – hydrophobic patches mediating ZPLD1 polymerization

After setting up the model system and providing proof of ZPLD1´s polymerization ability, I wanted to identify and initially describe the function of two hydrophobic patches involved in the regulation of ZPLD1 module-mediated interaction and polymerization. The location of these patches, an internal hydrophobic patch (IHP) as part of the bi-

partite ZP module and an extracellular hydrophobic patch (EHP) located distinct at the C-terminus is highly conserved between ZP proteins in terms of secondary structure as well as amino acid properties (Schaeffer et al., 2009). Based on these characteristics I identified both motifs in the zebrafish ZPLD1 sequence.

To understand if the regulated polymerization follows a universal principle, I analyzed phenotypic characteristics induced by mutations deleting the EHP and IHP of ZPLD1 using the established cellular model of MDCK cells. The observed results for ZPLD1 variants lacking one of the hydrophobic areas mirrors almost exactly prior results of Uromodulin variants (Schaeffer et al., 2009). The presence of both hydrophobic patches is required for proper polymerization, which strongly supports the hypothesis of their regulatory function. Already for Uromodulin the interesting additional examination has been reported that variants lacking the EHP traffic to the plasma membrane but seem to stick there without further processing. Interaction of the hydrophobic areas seems to be a prerequisite for the following proteolytic cleavage. A special conformational interaction between the two hydrophobic patches could expose the recognition site (CCS) of ZPLD1 for a so far unidentified protease or an environment that would activate cleavage (Schaeffer et al., 2009). Unfortunately, I was not yet able to confidently investigate conditioned medium and cell lysate via Western blot analysis to further justify the lack of secretion of the variant depleted of its putative EHP. I therefore tested this hypothesis by insertion and detection of additional His-tags shortly before and after the identified putative EHP regions. The results indicate that whilst the EHP seems to be important to enable proteolytic cleavage, is the sequence is lost upon ZPLD1 assembly. Obviously, the release of the EHP allows for proper extracellular protein polymerization into filaments. This step seems to be a universally valid as it even occurs in soluble ZP module proteins. Prominent examples are some fish vitelline envelope proteins. Secreted as precursors by liver cells the pro-proteins undergo proteolytic cleavage including loss of the EHP upon their arrival at the egg where they subsequently assemble into the egg coat (Darie et al., 2005, 2004).

4.3 The VIP: very important protease – ZPLD1 processing at the plasma membrane

Application of different protease inhibitors to transfected MDCK cells expressing ZPLD1 provided evidence that maturation of ZPLD1 requires a serine type protease. Similar results have already been shown for Uromodulin where the serine type protease Hepsin has been identified as responsible for its physiological release and subsequent filament formation (Brunati et al., 2015). Hepsin is a cell membrane-bound protease expressed in many tissues and of great importance in the inner ear (Brunati et al., 2015; Li et al., 2021). It is for example expressed in spiral ganglion neurons that innervate the hair cells of the inner ear (Guipponi et al., 2008, 2007). Interestingly, Hepsin KO mice have impaired hearing, among others due to defective cochlear structures with the exact underlying mechanism remaining unclear (Li et al., 2021). Together with all other reported findings of similarity between ZPLD1 and Uromodulin, Hepsin depicts a key potential candidate for the processing of ZPLD1 in the MDCK cell model as well as *in vivo*. However, Hepsin KO mice did not show cycling or other behavior hinting towards balance impairments (Guipponi et al., 2007). Consequently, further investigations, e.g., in accordance with the studies by Brunati and colleagues for Uromodulin, remain open to potentially confirm this hypothesis (Brunati et al., 2015).

At the same time inevitable, I confirmed ZPLD1´s cleavage site at position $_{315}$RKKR$_{318}$ using a corresponding depletion variant. The Immunofluorescence analyses of this variant again illustrated the implicit requirement of cleavage of the pro-peptide, by whichever protease, to allow for the polymerization of the mature ZP module.

4.4 Where the folding joins in – Structural considerations on ZPLD1 function and processing

Cleavage of the C-terminal pro-peptide might be crucial, as it includes the EHP. Most likely it needs to be removed as interplay of the hydrophobic patches avoids intracellular polymerization and shields the protein for intermolecular self-assembly as long, as it is inside the cell.

Structural findings for other ZP proteins based on X-ray and Cryo-EM studies further support this hypothesis. Results indicate that conformational changes and domain swapping events likely play a major role in ZP protein maturation and polymerization (Bokhove et al., 2016; Han et al., 2010; Stsiapanava et al., 2020). The EHP of ZP3 has for example been shown to block premature protein polymerization by acting as a "molecular glue" that keeps the ZP module in a conformation essential for secretion while hindering the formation of higher-order structures (Han et al., 2010). Assessment of AlphaFold structure prediction of ZPLD1 however, does not suggest an intramolecular interaction of its hydrophobic patches (Figure 12). While IHP and EHP do not seem to interact within a single ZPLD1 molecule their exposed location and orientation could instead enable homodimerization. Studies on crystal structures of Uromodulin already revealed homodimerization mediated by its hydrophobic patches as prerequisite for secretion, processing, and subsequent filament formation (Bokhove et al., 2016). In the same work Bokhove and colleagues hypothesize that numerous hydrophobic regions of Uromodulin interact in varying ways to enable conformations and folding as needed for secretion, processing, pre-dimerization, and polymerization. Thereby, they are weakening the hypothesis of IHP and EHP being the single or at least most important interacting switches for ZP module regulation (Bokhove et al., 2016). Based on these findings, the lacking proximity of the sequence areas in the (solely computational!) structural model does not exclude an important role for ZPLD1 functioning.

Localization of ZPLD1´s TMD and CCS using the AlphaFold model supports the reported immunofluorescence results. The TMD consists of an alpha-helix close to ZPLD1´s C terminus and being free to allow the idea of membrane association. The CCS is nicely exposed at the surface of the protein and seems easily accessible. Its backward orientation directed toward the protein´s C-terminus further indicates a certain proximity to the membrane surface, allowing cleavage by a membrane bound protease.

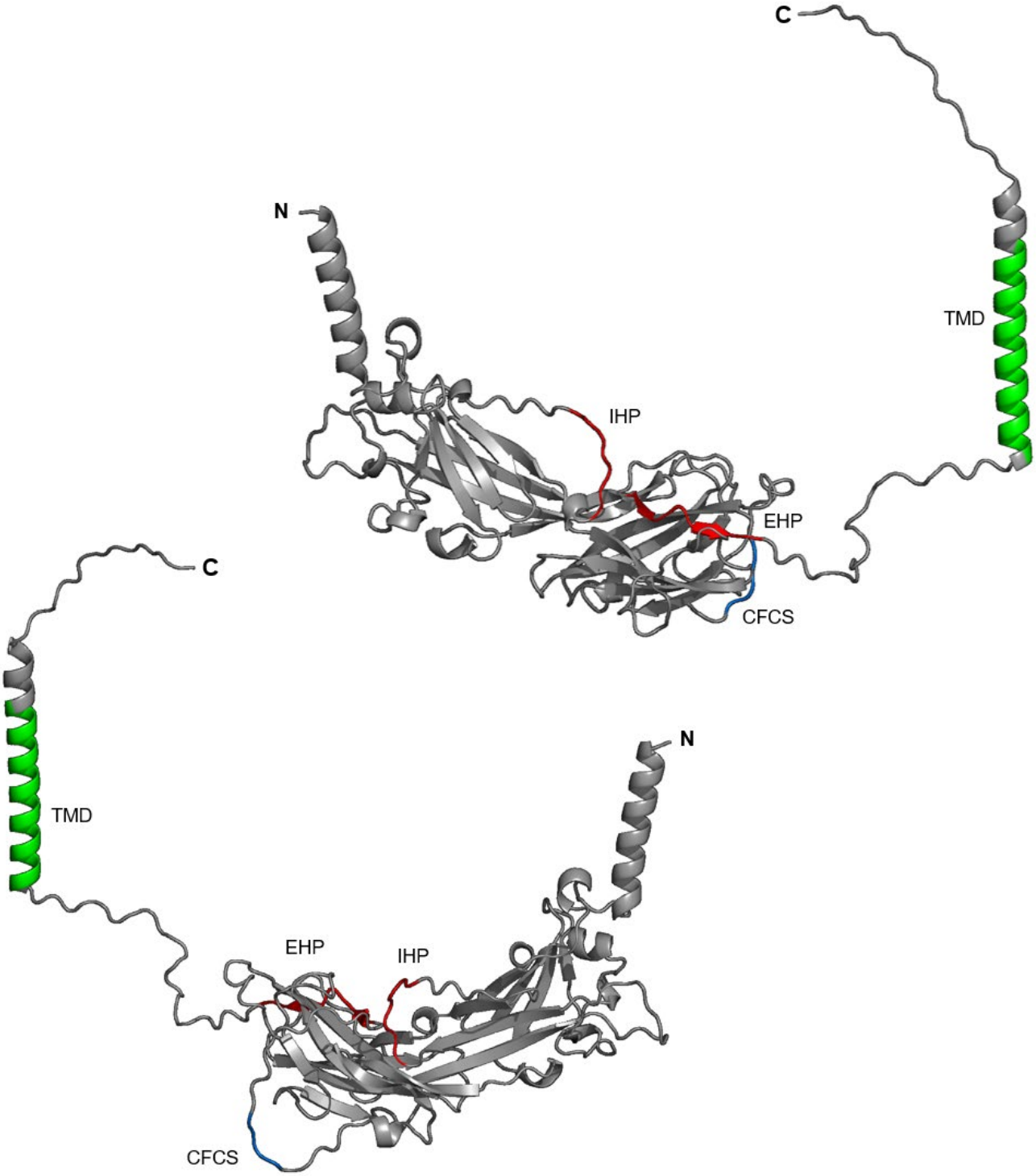


Figure 12. Visualization of the predicted structure of zebrafish ZPLD1 (accession A0A0R4ITH5) in two different orientations. Model prediction by AlphaFold (AlphaFold DB version 2022-06-01) ("AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models | Nucleic Acids Research | Oxford Academic," n.d.; Jumper et al., 2021). Putative IHP and EHP are marked in red. Putative TMD marked in green. Putative cleavage site marked in blue.

4.5 More to find in this region – Further interesting hints on the role of ZPLD1´s linker sequence

Another possible explanation for the lacking proximity of the identified hydrophobic patches in the structural model becomes evident when comparing the interdomain linker sequences of ZPLD1 and Uromodulin (Figure 13). Based on performed X-ray crystallographic studies of the polymerization region, Bokhove and colleagues described a rigid interdomain linker consisting of an alpha helix (α1), a beta strand (β1) and the consecutive IHP (Bokhove et al., 2016). The linker region has been shown to be essential for proper Uromodulin functioning, unlike ZP2 which is notably not able to form homopolymers (Bokhove et al., 2016). For Uromodulin, α1 packs tightly against the IHP and shields it from other hydrophobic regions, disabling a too tight connection between ZP-C and ZP-N (Bokhove et al., 2016). As a consequence, ZP-N is free to dimerize (Bokhove et al., 2016). This dimerization is an essential prerequisite for proper Uromodulin function as shown by mutation studies of involved residues completely abolishing filament formation (Bokhove et al., 2016). In summary, the rigid interdomain linker accounts for a polymerization competent conformation thereby enabling the formation of a Uromodulin homodimer which represents a polymerization intermediate, essential for proper polymerization (Bokhove et al., 2016).

This described organized ZP-N/ZP-C linker comprising of α1, β1 and IHP has already been shown to be conserved between self-polymerizing ZP modules such as glycoprotein 2 and the homopolymeric glycoprotein TECTA (Bokhove et al., 2016). TECTA can be found in the inner ear forming the tectorial membrane, an apical extracellular matrix that encloses the sensory hair cells required for hearing. The striking resemblance in location and function of TECTA and ZPLD1 serves as another hint towards a rigid linker structure and polymerization mechanism as described by Bokhove and colleagues (Bokhove et al., 2016).

Consecutive sequence alignment reveals that Schaeffer´s proposed model for the prediction of the IHP as well as the relevant hydrophobic sequence I identified for ZPLD1 might correspond to the α1 rather than the internal hydrophobic patch. Consensus of the linker structure seems furthermore likely as linker arrangements have been shown

to coincide with the different polymerization abilities of the corresponding ZP proteins (Bokhove et al., 2016).

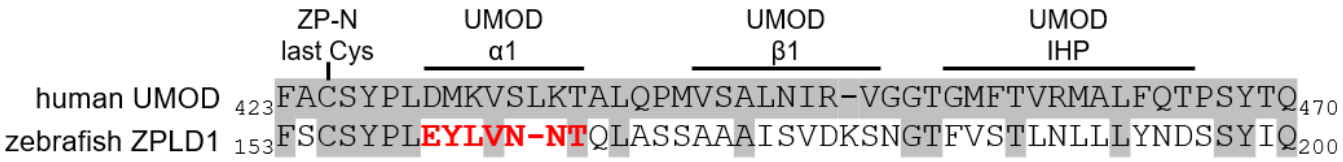


Figure 13. Sequence comparison of the interdomain linker sequences of zebrafish ZPLD1 (accession A0A0R4ITH5) and human Uromodulin (accession P07911). The structured rigid linker region of Uromodulin consists of an α-helix (α1), a beta-strand (β1) and the consecutive IHP (Bokhove et al., 2016). Alignment reveals that Schaeffer´s proposed model for the prediction of the IHP as well as the relevant hydrophobic sequence we identified for ZPLD1 (marked in red) might correspond to the α1-helix rather than the internal hydrophobic patch.

However, the secondary structure predictions are inconsistent. AlphaFold as well as four out of five algorithms used for the structural predictions to identify the putative IHP do not anticipate a helix for the identified hydrophobic sequence region (Figure 5). This fact underlines uncertainties posed by the sole availability of computational structure predictions. Further investigation will be needed to provide more evidence and details to the structural arrangement of the interdomain linker of ZPLD1 and its role in the regulated polymerization process, which is likely common for many, if not all of ZP-proteins.

4.6 Putting it all together – an initial model of ZPLD1 polymerization

The presented work investigated for the first-time polymerization and underlying mechanisms of ZPLD1. I demonstrated that short hydrophobic motifs common for ZP proteins could also be detected in ZPLD1 using proposed conserved amino acid features. These regions play a similar role for regulation and polymerization as for the classical ZP module proteins. Furthermore, I was able to show the requirement of a cleavage site as well as membrane anchoring via a transmembrane domain to enable maturation and secretion of ZPLD1 facilitated by a membrane bound serine type protease.

The mechanism widely mirrors earlier findings for homodimerizing ZP-proteins. Consequently, I hypothesize a model for ZPLD1 regulation and processing that is likely conserved between this class of ZP family proteins (Figure 14). Interaction between the hydrophobic patches in the linker as well as the pro-peptide results in a conformational state, which enables proper attachment to the vesicular membrane during transport to the cytoplasmic membrane (CM). Anchoring of immature ZPLD1 to the CM is a prerequisite for further maturation. Interplay of the hydrophobic patches avoids intracellular polymerization and shields the protein for intermolecular self-assembly. Cleavage of the EHP as part of the C-terminal pro-peptide, by a serine type protease, potentially Hepsin, results in the release of mature ZPLD1 and paves the way for extracellular dimerization. ZP-N mediated dimers finally serve as building blocks for further assembly into higher molecular structures.

This model together with successful experiments of other well studied ZP proteins could well provide a blueprint to further study function, polymerization mechanism and regulation of ZPLD1.

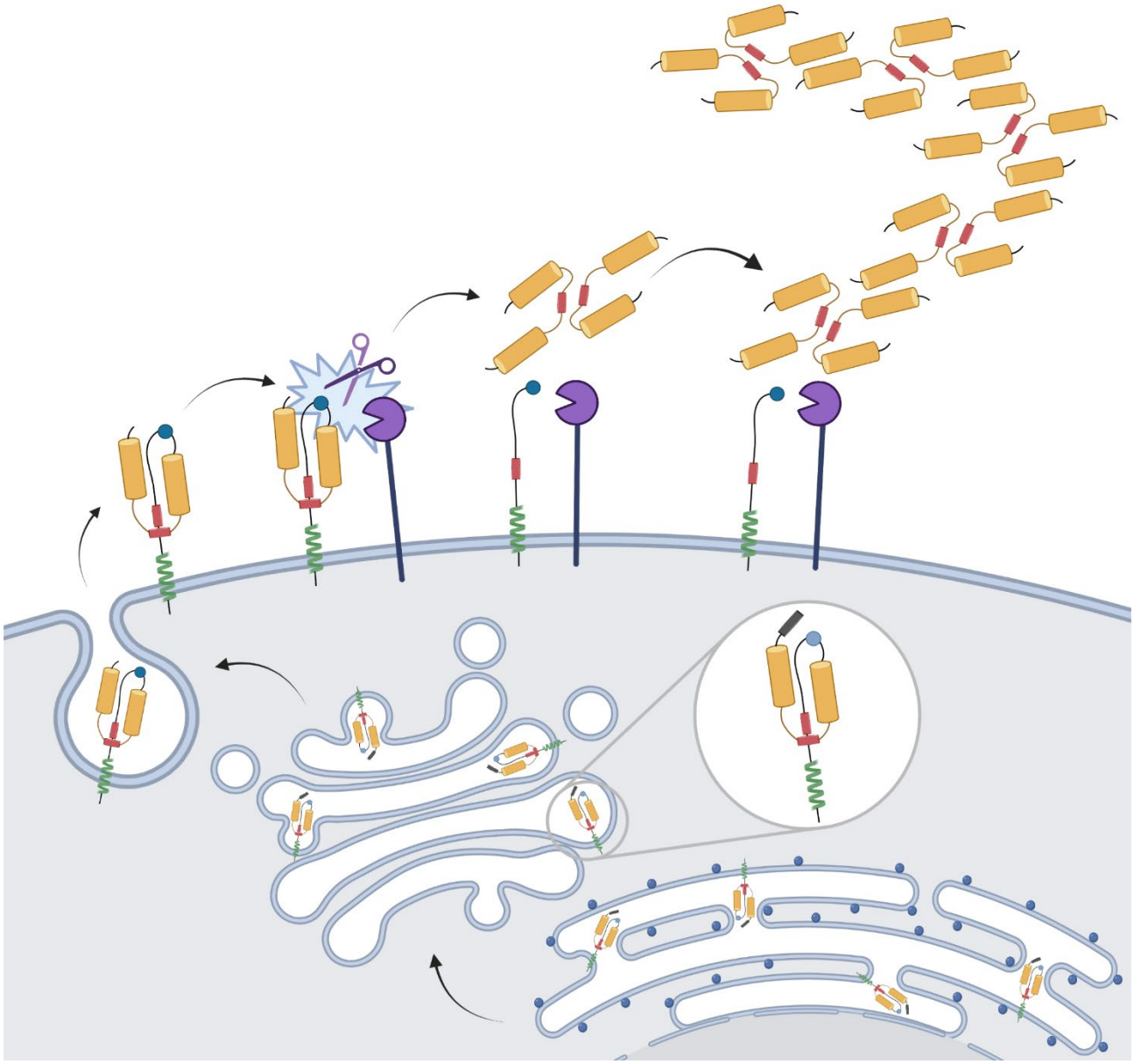


Figure 14. Model of ZPLD1 shedding and polymerization. The protein is expressed, enters the secretory pathway led by its signal peptide (black cylinder), and reaches the plasma membrane in a polymerization-incompetent confirmation. This is ensured by the interaction between the hydrophobic patches in the linker as well as the pro-peptide (short red cylinders). Processing by a serine type protease (purple) likely occurs at the plasma membrane. Shedding at the cleavage site (blue dot) releases the hydrophobic interaction of the patches, generating polymerization-competent species that assemble first into dimeric building blocks and subsequently into higher molecular polymeric filaments to establish the cupula structure. Yellow cylinders represent ZP-N and ZP-C subdomains, the transmembrane domain is shown in green. Figure created with BioRender.com ("Scientific Image and Illustration Software | BioRender," n.d.).

4.7 Cupula hydrogel and beyond – ZPLD1 remains an interesting research target

Based on its ability to polymerize, ZPLD1 likely presents the structural component ensuring correct formation of the cupula in the inner ear. Thereby, it serves as an important factor for proper vestibular function. The fact that ZPLD1 is part of the group of homodimerizing ZP proteins which seem to share a magnitude of regulatory mechanisms and characteristics provides and additional indication for the importance of this type of proteins for living and biological species in general. In accordance, many pathologies are related to (genetic mutations) resulting in ZP protein dysfunction. Interestingly the ZP module seems to be a universal tool for the formation of hydrogels and polymeric structures in a variety of different settings. Further investigations on ZP proteins, their functioning, and possibilities to compensate for damages or losses therefore appear worthy. Bionic usage of ZP module assembly e.g., for the synthesis of artificial hydrogels might as well pose an attractive research target.

Another interestingly aspect for ZPLD1 in particular is the variety of pathologies where genome and/or transcriptome wide analyses hint towards an involvement of the ZP protein. ZPLD1´s accumulating and homopolymerizing properties would fit well to underlying hypotheses around possible involvement in cell-cell interaction, adhesion, and migration. Based on these findings cell type interaction studies, ZPLD1 knock-out studies as well as systematic transcriptome and expression level analyses might be interesting to gain further insights in ZPLD1´s roles and involvements.

Taking all aspects together, I strongly believe ZPLD1 is an interesting target for further investigations as it seems not solely of great relevance for proper vestibular function but in addition to further adhesive biological functions.

5 Additional materials and methods

5.1 Additional ZPLD1 expression constructs

Additional genetic manipulations (deletions and alterations) were as well performed using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) in combination with back-to-back primer design (see details in materials and methods section above). Primers were ordered from Metabion international AG. Primer sequences for the additional ZPLD1 expression constructs can be found in Table 2.

Target sequences for the deletion of the putative β1 as well as the second (new) putative IHP region were identified performing sequence alignments of ZPLD1 and Uromodulin (see Discussion section 4.5, Figure 13).

Position and likelihood of asparagine residues of potential *N*-glycosylation sites (NXS/T) were predicted using the program NetNGlyc 1.0 (Center for Biological Sequence Analysis, Technical University of Denmark) as previously described (Dernedde et al., 2014; Gupta and Brunak, 2002). The four altered *N*-glycosylation sites (according to their calculated likelihoods) N_{121}, N_{164}, N_{181} and N_{194}, were all mutated to a glutamine (Q).

All constructs were sequence-verified before transfection.

Table 2: Primers (5´-3´) used to generate additional ZPLD1 expression constructs

Mutation	Primer
Deletion of the putative β1 region	Forward: GGTACTTTTGTTAGCACGCTCAATTTGC
	Reverse: CGACGAAGCAAGCTGTGTGTTGTTG
Deletion of the second putative IHP region (ΔIHP II)	Forward: TCCTCATACATTCAGAACCTCGCC
	Reverse: AGTACCATTACTGTCCTTCACTGATATCGC
Alteration of N_{121}	Forward: CCCAACGCCTATGGGACCTTTCCCTGGTCC
	Reverse: CCCATGAGCCGTAGAGACCACC
Alteration of N_{164}	Forward: GGAATACCTGGTCCAGAACACACAGCTTGC
	Reverse: AGTGGGTAGCTACAACTGAACTTGTAAAGC
Alteration of N_{181}	Forward: TGAAGGACAGTCAAGGTACTTTTGTTAGC
	Reverse: CTGATATCGCTGCGGCCGACGAAGCAAGc
Alteration of N_{194}	Forward: CAATTTGCTTCTATACCAGGACTCCTCATAc
	Reverse: AGCGTGCTAACAAAAGTACCATTACTGTCC

5.2 Native PAGE analysis

Native gradient gels were prepared ranging from 4% to 20% without addition of any SDS or other reducing agents. Purified and concentrated ZPLD1 as well as used standard protein ladder samples were prepared in native sample buffer containing glycerin and bromophenol blue. Native gels were run at 100 V and 4°C for around 15 h. Subsequently, gels were cut to allow for Coomassie blue staining of lanes containing protein standard as well as Western blot analysis of ZPLD1 (see Materials and Methods section 2.4)

5.3 Electron microscopy

For electron microscopy (EM) ZPLD1 was purified and concentrated as described above to a final concentration of 1 mg/ml. Preparation for transmission electron microscopy (TEM) and final TEM imaging were kindly performed by Dr. Kai Ludwig at the BioSupraMol of the Freie Universität Berlin as follows. A 5 µl droplet of the ZPLD1 sample solution was applied to a hydrophilized Formvar-carbon film covering an EM copper grid (400 mesh). The supernatant liquid was removed by blotting with a filter paper, and the sample was allowed to air dry. A contrast-enhancing heavy metal staining solution (1% uranyl acetate) was then applied for 60 s and blotted again ("negative staining method"). The dried samples were transferred with a standard holder into a Talos L120C transmission electron microscope (Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA) with a tungsten cathode at an accelerating voltage of 120 kV. Microscopic images were acquired using a Ceta 16M camera.

6 Additional preliminary results

6.1 *N*-glycosylation of ZPLD1 might be essential for protein function

The established expression model was used to study ZPLD1´s *N*-glycosylation by creation and initial analysis of four additional altered ZPLD1 variants. ZPLD1´s observed molecular weight after PNGase-F deglycosylation is reduced by ~11 kDa (Figure 4) (Dernedde et al., 2014). Based on these findings, a posttranslational modification by three to four *N*-glycan chains can be expected, depending on their individual structure (Dernedde et al., 2014). Most likely potential *N*-glycosylation sites have been predicted using a computational program (Dernedde et al., 2014). Accordingly, four separate expression constructs each containing a mutation for one of the potential sites were created.

Initial experiments interestingly reveal, when transfected to MDCK cells, all four ZPLD1 variants seem to lack their ability to properly form polymeric structures on the outer cell membrane (Figure 15 A). Failure to detect polymers due to lacking expression or low expression levels could be excluded based on the assessment of permeabilized cells (Figure 15 B). Some areas of pictures taken from permeabilized cells might even suggest early-stage formation of internal polymers or aggregation.

The observed initial results match the fact that, polymerization defects were also reported for a mutation in a *N*-glycosylation site of TECTA (Sagong et al., 2010). This mutation is associated with hearing loss suggesting importance of this carbohydrate for tectorial membrane assembly and has been shown to be conserved among filament forming ZP modules (Bokhove et al., 2016; Sagong et al., 2010). Correspondingly, a glycan within the β-sheet region inside the ZP-N domain of Uromodulin has been described to be essential for proper polymerization functioning (Bokhove et al., 2016). It forms intermolecular hydrogen bonds with a ZP-N domain of another Uromodulin molecule thereby enabling formation of the dimeric polymerization intermediate (Bokhove et al., 2016).

Based on these insights and the shown initial findings, *N*-glycosylation of ZPLD1 could as well play a relevant, if yet unclear role for functioning or regulation, making further investigations worthy.

6.2 Further sequence motifs within ZPLD1´s linker region could be involved in the polymerization process

As discussed above, sequence comparisons of the interdomain linker sequences of human uromodulin and zebrafish ZPLD1 indicate that the prediction of the IHP as well as the corresponding hydrophobic sequence identified for ZPLD1 might correspond to the α1-helix rather than the internal hydrophobic patch (Figure 13). Consecutive, the sequences region $_{172}$AAAISVDKS$_{180}$ might correspond with ZPLD1´s β1-strand and $_{184}$FVSTLNLLLYND$_{195}$ with its actual IHP. For an initial test of this new hypothesis, expression constructs for ZPLD1 variants lacking the corresponding sequence regions were prepared and transfected to MDCK cells (Figure 16 A).

Immunofluorescence analysis reveals, that both regions are indeed as well of importance for proper ZPLD1 polymerization (Figure 16 B). Interestingly, observation of unpermeabilized cells transfected with the variant lacking the second putative IHP region (ΔIHP II) results in the same outcome/phenotype as for the initial putative region (ΔIHP, Figure 6). It seems like early forms of polymers start to aggregate but stay abnormally short and less organized. When investigating permeabilized cells however, intracellular polymers are not as clearly recognized as compared to the ΔIHP variant (Figure 16 C). Samples of unpermeabilized cells transfected with the Δβ1 variant do not show any polymer formation at all (Figure 16 B). Compared to the other tested linker variants lower amounts of protein are observed on the outer cell membrane. This may be due to faster processing and cut-off through the responsible proteases. Another reason could be lower expression levels or intracellular degradation, as permeabilized cells seem as well to provide a lower signal for anti- strep-tag positive protein (Figure 16 C). However, shown results are only preliminary, as experiments have to the current state, not been sufficiently replicated.

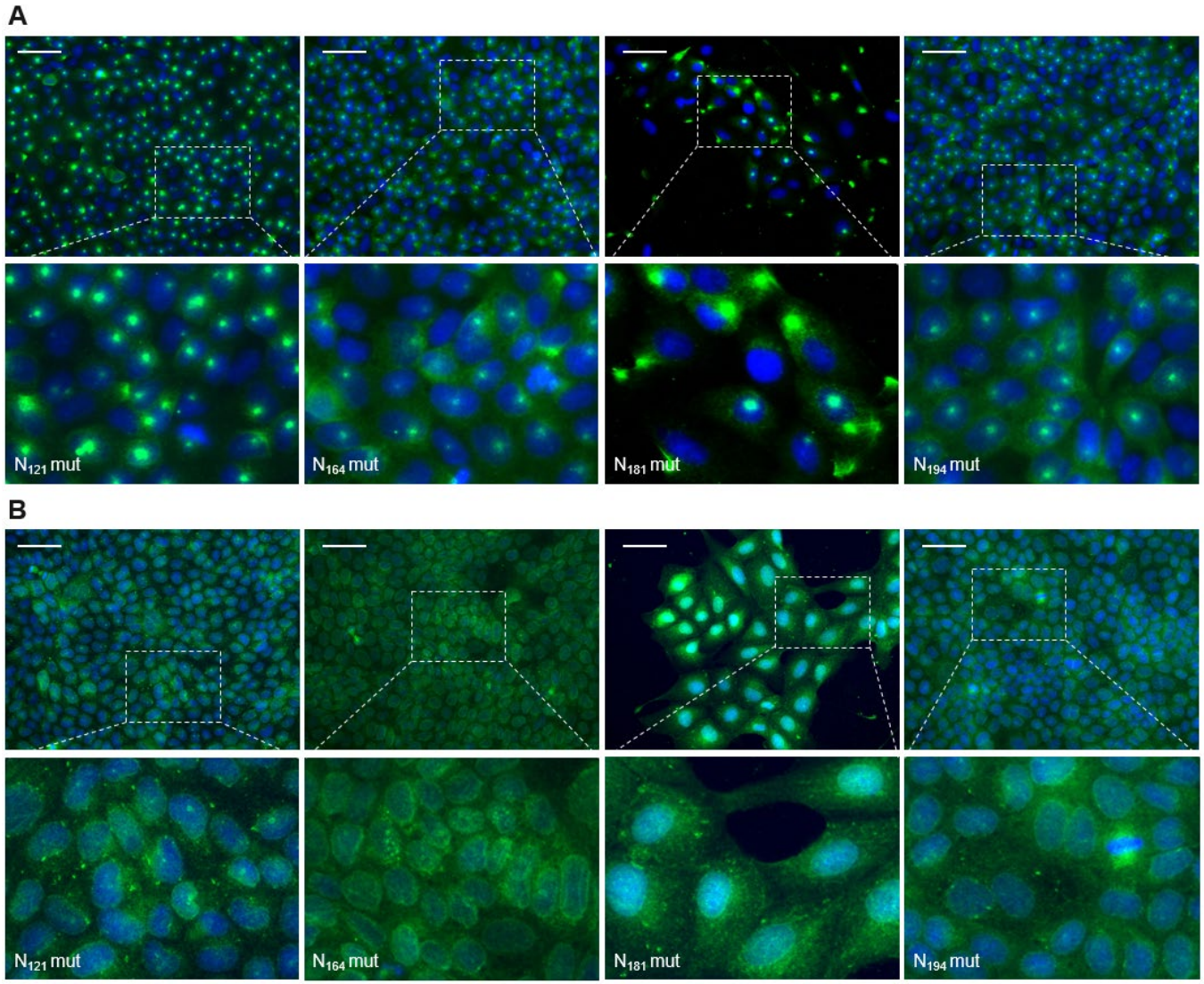


Figure 15: Alteration of every one of the most probable *N*-glycosylation sites results in impaired ability of ZPLD1 to form polymers on the surface of transfected MDCK cells. Immunofluorescence of MDCK cells stably expressing ZPLD1 N_{121}, N_{164}, N_{181}, and N_{194} mutant variants. (A) non-permeabilized and (B) permeabilized MDCK cells. The presented merged pictures show Strep (ZPLD1) in green; DAPI cell nucleus staining is shown in blue. Bar, 50 µm. Preliminary results are shown. To this date, experiments were not yet properly replicated.

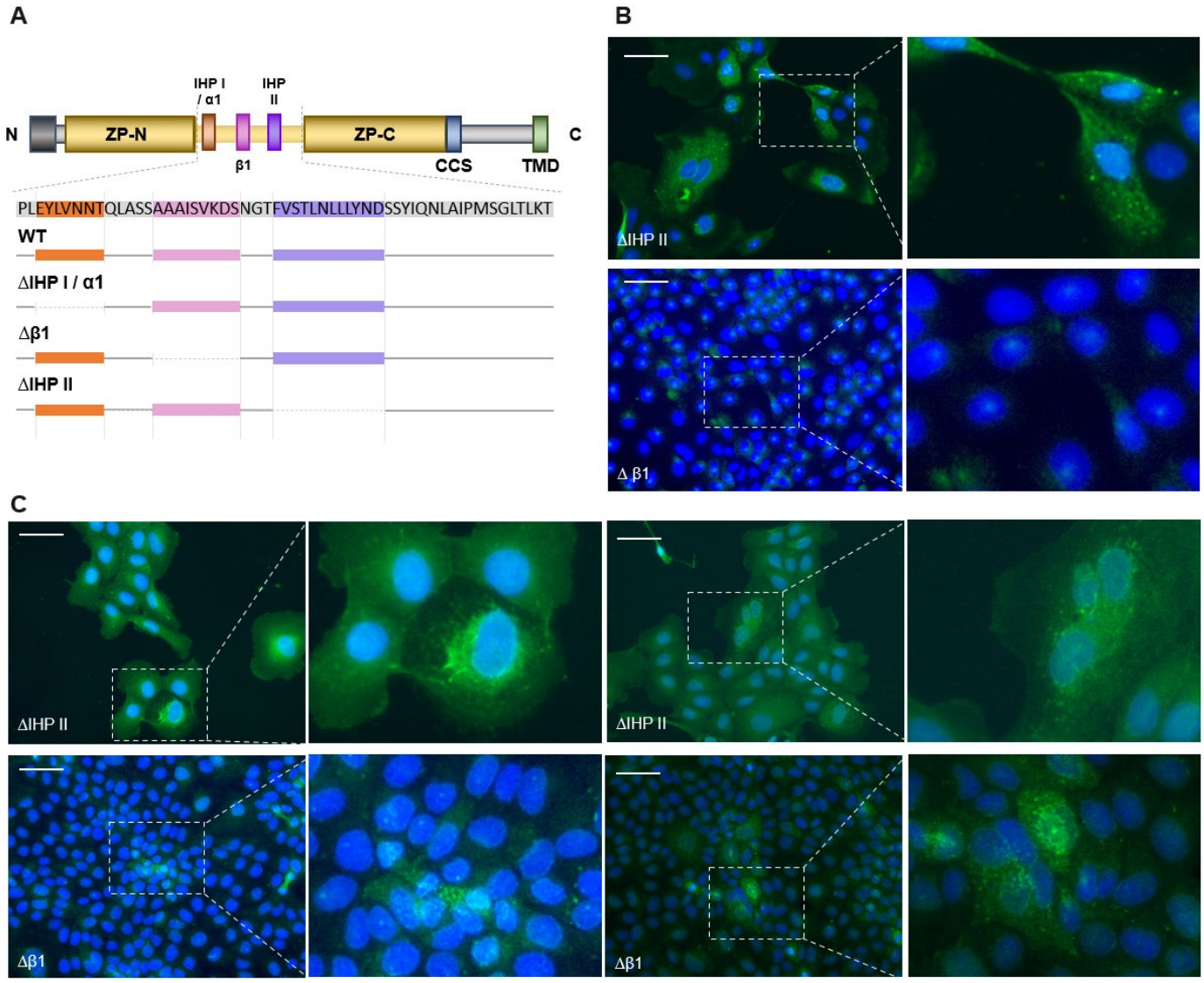


Figure 16. Deletion of the β1 sequence or second (new) putative IHP regions result in hindered polymerization and intracellular structures in the broadest sense comparable to the intracellular polymers observed in EHP and IHP depleted variants. (A) Domain structure and sequence excerpt of ZPLD1. The putative IHP I / α1 is highlighted in orange, β1 is highlighted in pink and IHP II highlighted in purple. All marked sequence areas were completely deleted in the respective Δ variants. Immunofluorescence of MDCK cells stably expressing ZPLD1 β1 and second (new) IHP mutant proteins. (B) non-permeabilized and (C) permeabilized MDCK cells. The presented merged pictures show Strep (ZPLD1) and in green, DAPI cell nucleus staining is shown in blue. Bar, 50 µm. Preliminary results are shown. To this date, experiments were not yet properly replicated.

To gain deeper insights in structure and function of ZPLD1´s interdomain linker further experiments will be needed. First, observation of variants wearing mutated instead of deleted target sequences will need to exclude steric effects based on a shorter linker peptide. Ultimately, assessment of crystal structures of ZPLD1, as performed e.g., for Uromodulin, are needed to confirm hypothesis regarding similarities between those two and other ZP proteins (Bokhove et al., 2016).

6.3 ZPLD1 seems to form high molecular weight complexes assembling dimeric building blocks

Initial analysis of purified ZPLD1 using native PAGE was performed to gain first insights into polymer sizes and ratios. ZPLD1 forms high molecular weight polymers of more than 450 kDa (Figure 17). Obviously, this accounts for more than 10 single ZPLD1 proteins which can be involved in these complexes. Interestingly, complexes below ~180 kDa cannot be observed on the Western blot. A possible explanation could be the relatively lower number of epitopes accessible for the primary antibody.

Another explanation could be as simple as timing. It could be that the portion of low molecular weight polymers decreases over time as higher molecular structures form. However, this would require mature ZPLD1 to be able to interact rather spontaneously upon steric proximity. Findings on Uromodulin and ZP3 rather support a certain mechanism requiring membrane association to enable polymerization (Bokhove et al., 2016; Han et al., 2010; Stanisich et al., 2020). A series of native PAGE experiments from a single sample over time could be used to initially test this hypothesis.

Furthermore, the observation could potentially indicate, that dimeric or tetrameric ZPLD1 “building blocks” are initially formed before being assembled into even higher molecular weight structures. This hypothesis is supported by the observed ladder-like pattern of polymers on the blot.

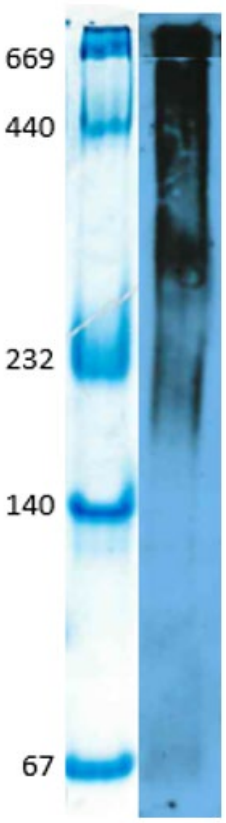


Figure 17. Native PAGE analysis of purified ZPLD1 reveals the formation of high molecular weight polymers containing more than ten single proteins. Small complexes containing less than four single proteins cannot be observed. This finding, together with the overall ladder-pattern, suggests the assembly of potentially dimeric precursors or "building blocks", subsequently incorporated into even higher molecular structures. Depicted numbers represent kDa weight of the Coomassie-blue stained protein standard components. Preliminary results are shown. To this date, experiments were not yet properly replicated.

The existence of dimeric precursors has been already described for other ZP-proteins such as Uromodulin and ZP3 (Bokhove et al., 2016; Han et al., 2010). To visualize polymers, ZPLD1 purified from MDCK supernatants has been further analyzed in first attempts using electron microscopy performed by Dr. Kai Ludwig from the BioSupraMol core facility of the Freie Universität Berlin. The first overview images at low magnification (Figure 18) show (supposedly positively contrasted) structures resembling ZPLD1 polymers as observed in the immunofluorescence analyses (Figure 4 B). At higher magnification, numerous negatively contrasted structures are visible, which appear to be particularly densely associated in the region of the apparent filaments, but no concrete conclusions can be drawn about dimer- or tetramerization as part of the

polymerization mechanism, the formation of "building blocks" or the polymer structure (Figure 18 B and C).

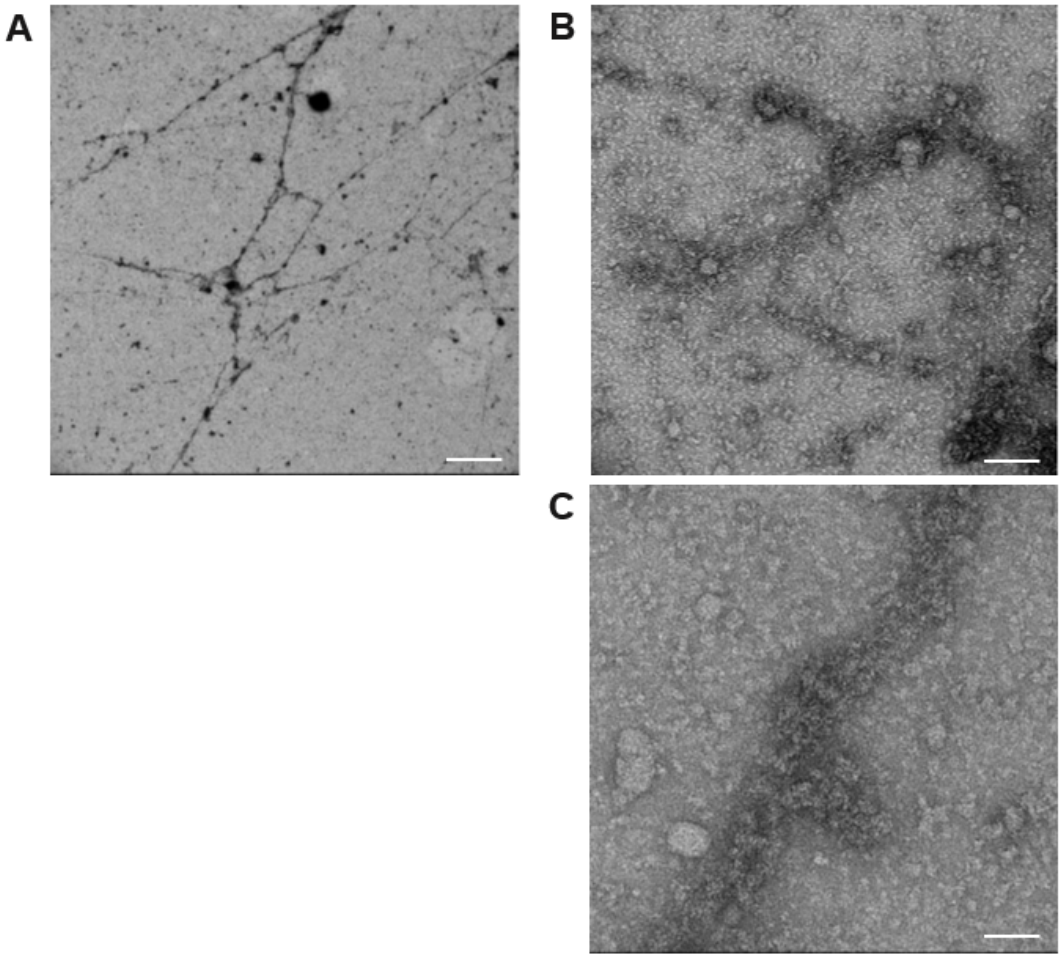


Figure 18. Electron microscopy analyses of negatively stained, purified ZPLD1 samples show structures which resemble polymers as observed in immunofluorescence pictures but lack resolution to state clear conclusions regarding repeating units or polymerization mechanism (A) Bar, 2 µm. (B) Bar, 100 nm. (C) Bar, 50 nm.

To improve quality and informative value, impurities need to be minimized and related visual distortions ruled out. A staining with gold labelled antibodies could be used to proof visualization of ZPLD1 and provide insight regarding potential repeating units as observed e.g., in electron micrographs of gold labelled ZP3 mutants (Jovine et al., 2006) Unfortunately, I was not yet able to purify ZPLD1 in sufficient quantity and concentration to conduct required experiment replications and optimize the setup accordingly. The same is true for cryo-EM studies, which could provide high resolution under native

conditions but require a particularly pure and sufficiently highly concentrated protein sample. Above all, future X-ray crystallographic studies will be needed to complete insights on ZPLD1´s polymerization mechanism (Bokhove et al., 2016).

List of abbreviations

CCM	Cerebral cavernous malformations
CM	Cytoplasmic membrane
CCS	Conserved consensus cleavage site
DAPI	4,6-diamidino-2-phenylindole
EHP	External hydrophobic patch
EM	Electron microscopy
IgG	Immunoglobulin G
IHP	Internal hydrophobic patch
MDCK	Mardin-Darby canine kidney
PFA	Paraformaldehyde
PIC	Protease inhibitor cocktail
SP	Signal peptide
TECTA	α-tectorin
TEM	Transmission electron microscopy
TMD	Transmembrane domain
ZP	Zona Pellucida
ZP-C	ZP-C domain (C-terminal subdomain of the ZP module)
ZP-N	ZP-N domain (N-terminal subdomain of the ZP module)
ZPLD1	Zona Pellucida Like Domain 1 Protein

References

AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models | Nucleic Acids Research | Oxford Academic [WWW Document], n.d. URL https://academic.oup.com/nar/article/50/D1/D439/6430488 (accessed 10.14.22).

Bokhove, M., Jovine, L., 2018. Structure of Zona Pellucida Module Proteins. Curr Top Dev Biol 130, 413–442. https://doi.org/10.1016/bs.ctdb.2018.02.007

Bokhove, M., Nishimura, K., Brunati, M., Han, L., de Sanctis, D., Rampoldi, L., Jovine, L., 2016. A structured interdomain linker directs self-polymerization of human uromodulin. Proc Natl Acad Sci U S A 113, 1552–1557. https://doi.org/10.1073/pnas.1519803113

Bork, P., Sander, C., 1992. A large domain common to sperm receptors (Zp2 and Zp3) and TGF-beta type III receptor. FEBS Lett 300, 237–240. https://doi.org/10.1016/0014-5793(92)80853-9

Brunati, M., Perucca, S., Han, L., Cattaneo, A., Consolato, F., Andolfo, A., Schaeffer, C., Olinger, E., Peng, J., Santambrogio, S., Perrier, R., Li, S., Bokhove, M., Bachi, A., Hummler, E., Devuyst, O., Wu, Q., Jovine, L., Rampoldi, L., 2015. The serine protease hepsin mediates urinary secretion and polymerisation of Zona Pellucida domain protein uromodulin. Elife 4, e08887. https://doi.org/10.7554/eLife.08887

Chen, Y.-C., Gonzalez, M.E., Burman, B., Zhao, X., Anwar, T., Tran, M., Medhora, N., Hiziroglu, A.B., Lee, W., Cheng, Y.-H., Choi, Y., Yoon, E., Kleer, C.G., 2019. Mesenchymal Stem/Stromal Cell Engulfment Reveals Metastatic Advantage in Breast Cancer. Cell Rep 27, 3916-3926.e5. https://doi.org/10.1016/j.celrep.2019.05.084

Cheng, J., Randall, A.Z., Sweredoski, M.J., Baldi, P., 2005. SCRATCH: a protein structure and structural feature prediction server. Nucleic Acids Res 33, W72-76. https://doi.org/10.1093/nar/gki396

Darie, C.C., Biniossek, M.L., Gawinowicz, M.A., Milgrom, Y., Thumfart, J.O., Jovine, L., Litscher, E.S., Wassarman, P.M., 2005. Mass spectrometric evidence that proteolytic processing of rainbow trout egg vitelline envelope proteins takes place on the egg. J Biol Chem 280, 37585–37598. https://doi.org/10.1074/jbc.M506709200

Darie, C.C., Biniossek, M.L., Jovine, L., Litscher, E.S., Wassarman, P.M., 2004. Structural characterization of fish egg vitelline envelope proteins by mass spectrometry. Biochemistry 43, 7459–7478. https://doi.org/10.1021/bi0495937

Dernedde, J., Weise, C., Müller, E.-C., Hagiwara, A., Bachmann, S., Suzuki, M., Reutter, W., Tauber, R., Scherer, H., 2014. Cupulin is a zona pellucida-like domain protein and major component of the cupula from the inner ear. PLoS One 9, e111917. https://doi.org/10.1371/journal.pone.0111917

Dohlman, G.F., 1969. The shape and function of the cupula. J Laryngol Otol 83, 43–53. https://doi.org/10.1017/s0022215100070031

Drozdetskiy, A., Cole, C., Procter, J., Barton, G.J., 2015. JPred4: a protein secondary structure prediction server. Nucleic Acids Res 43, W389-394. https://doi.org/10.1093/nar/gkv332

Gianfrancesco, F., Esposito, T., Penco, S., Maglione, V., Liquori, C.L., Patrosso, M.C., Zuffardi, O., Ciccodicola, A., Marchuk, D.A., Squitieri, F., 2008. ZPLD1 gene is disrupted in a patient with balanced translocation that exhibits cerebral cavernous malformations. Neuroscience 155, 345–349. https://doi.org/10.1016/j.neuroscience.2008.05.030

Glessner, J.T., Bradfield, J.P., Wang, K., Takahashi, N., Zhang, H., Sleiman, P.M., Mentch, F.D., Kim, C.E., Hou, C., Thomas, K.A., Garris, M.L., Deliard, S., Frackelton, E.C., Otieno, F.G., Zhao, J., Chiavacci, R.M., Li, M., Buxbaum, J.D., Berkowitz, R.I., Hakonarson, H., Grant, S.F.A., 2010. A genome-wide study reveals copy number variants exclusive to childhood obesity cases. Am J Hum Genet 87, 661–666. https://doi.org/10.1016/j.ajhg.2010.09.014

Guipponi, M., Tan, J., Cannon, P.Z.F., Donley, L., Crewther, P., Clarke, M., Wu, Q., Shepherd, R.K., Scott, H.S., 2007. Mice deficient for the type II transmembrane serine protease, TMPRSS1/hepsin, exhibit profound hearing loss. Am J Pathol 171, 608–616. https://doi.org/10.2353/ajpath.2007.070068

Guipponi, M., Toh, M.-Y., Tan, J., Park, D., Hanson, K., Ballana, E., Kwong, D., Cannon, P.Z.F., Wu, Q., Gout, A., Delorenzi, M., Speed, T.P., Smith, R.J.H., Dahl, H.H., Petersen, M., Teasdale, R.D., Estivill, X., Park, W.J., Scott, H.S., 2008. An integrated genetic and functional analysis of the role of type II transmembrane serine proteases (TMPRSSs) in hearing loss. Human Mutation 29, 130–141. https://doi.org/10.1002/humu.20617

Gupta, R., Brunak, S., 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput 310–322.

Han, L., Monné, M., Okumura, H., Schwend, T., Cherry, A.L., Flot, D., Matsuda, T., Jovine, L., 2010. Insights into Egg Coat Assembly and Egg-Sperm Interaction from the X-Ray Structure of Full-Length ZP3. Cell 143, 404–415. https://doi.org/10.1016/j.cell.2010.09.041

Helling, K., Clarke, A.H., Watanabe, N., Scherer, H., 2000. [Morphological studies of the form of the cupula in the semicircular canal ampulla]. HNO 48, 822–827. https://doi.org/10.1007/s001060050668

Iimura, Y., Suzuki, M., Otsuka, K., Inagaki, T., Konomi, U., Shimizu, S., 2010. Effect of cupula shrinkage on the semicircular canal activity. Acta Otolaryngol 130, 1092–1096. https://doi.org/10.3109/00016481003677456

ImageJ [WWW Document], n.d. URL https://imagej.net/ij/index.html (accessed 2.14.23).

Jovine, L., Darie, C.C., Litscher, E.S., Wassarman, P.M., 2005. Zona pellucida domain proteins. Annu Rev Biochem 74, 83–114. https://doi.org/10.1146/annurev.biochem.74.082803.133039

Jovine, L., Janssen, W.G., Litscher, E.S., Wassarman, P.M., 2006. The PLAC1-homology region of the ZP domain is sufficient for protein polymerisation. BMC Biochem 7, 11. https://doi.org/10.1186/1471-2091-7-11

Jovine, L., Qi, H., Williams, Z., Litscher, E.S., Wassarman, P.M., 2004. A duplicated motif controls assembly of zona pellucida domain proteins. Proc Natl Acad Sci U S A 101, 5922–5927. https://doi.org/10.1073/pnas.0401600101

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kobayashi, D., Nishizawa, D., Takasaki, Y., Kasai, S., Kakizawa, T., Ikeda, K., Fukuda, K., 2013. Genome-wide association study of sensory disturbances in the inferior alveolar nerve after bilateral sagittal split ramus osteotomy. Mol Pain 9, 34. https://doi.org/10.1186/1744-8069-9-34

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305, 567–580. https://doi.org/10.1006/jmbi.2000.4315

Li, S., Wang, L., Sun, S., Wu, Q., 2021. Hepsin: a multifunctional transmembrane serine protease in pathobiology. FEBS J 288, 5252–5264. https://doi.org/10.1111/febs.15663

Liu, X.-Q., Fazio, J., Hu, P., Paterson, A.D., 2016. Identity-by-descent mapping for diastolic blood pressure in unrelated Mexican Americans. BMC Proc 10, 263–267. https://doi.org/10.1186/s12919-016-0041-x

McGuffin, L.J., Bryson, K., Jones, D.T., 2000. The PSIPRED protein structure prediction server. Bioinformatics 16, 404–405. https://doi.org/10.1093/bioinformatics/16.4.404

Meng, Z., Yuan, Q., Zhao, J., Wang, B., Li, S., Offringa, R., Jin, X., Wu, H., 2020. The m6A-Related mRNA Signature Predicts the Prognosis of Pancreatic Cancer Patients. Mol Ther Oncolytics 17, 460–470. https://doi.org/10.1016/j.omto.2020.04.011

Pak, J., Pu, Y., Zhang, Z.-T., Hasty, D.L., Wu, X.-R., 2001. Tamm-Horsfall Protein Binds to Type 1 Fimbriated Escherichia coli and Prevents E. coli from Binding to Uroplakin Ia and Ib Receptors *. Journal of Biological Chemistry 276, 9924–9930. https://doi.org/10.1074/jbc.M008610200

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8, 785–786. https://doi.org/10.1038/nmeth.1701

Pollastri, G., McLysaght, A., 2005. Porter: a new, accurate server for protein secondary structure prediction. Bioinformatics 21, 1719–1720. https://doi.org/10.1093/bioinformatics/bti203

PyMOL | pymol.org [WWW Document], n.d. URL https://pymol.org/2/#page-top (accessed 10.20.22).

Rost, B., Liu, J., 2003. The PredictProtein server. Nucleic Acids Res 31, 3300–3304. https://doi.org/10.1093/nar/gkg508

Sagong, B., Park, R., Kim, Y.H., Lee, K.-Y., Baek, J.-I., Cho, H.-J., Cho, I.-J., Kim, U.-K., Lee, S.-H., 2010. Two novel missense mutations in the TECTA gene in Korean families with autosomal dominant nonsyndromic hearing loss. Ann Clin Lab Sci 40, 380–385.

Schaeffer, C., Santambrogio, S., Perucca, S., Casari, G., Rampoldi, L., 2009. Analysis of uromodulin polymerization provides new insights into the mechanisms regulating ZP domain-mediated protein assembly. Mol Biol Cell 20, 589–599. https://doi.org/10.1091/mbc.e08-08-0876

Scherer, H., Watanabe, S., 2001. Introductory remarks on this issue. On the role of the ampulla in disturbances of vestibular function. Biol Sci Space 15, 350–352. https://doi.org/10.2187/bss.15.350

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9, 671–675. https://doi.org/10.1038/nmeth.2089

Scientific Image and Illustration Software | BioRender [WWW Document], n.d. URL https://www.biorender.com/ (accessed 7.21.23).

Stanisich, J.J., Zyla, D.S., Afanasyev, P., Xu, J., Kipp, A., Olinger, E., Devuyst, O., Pilhofer, M., Boehringer, D., Glockshuber, R., 2020. The cryo-EM structure of the human uromodulin filament core reveals a unique assembly mechanism. Elife 9, e60265. https://doi.org/10.7554/eLife.60265

Stsiapanava, A., Xu, C., Brunati, M., Zamora-Caballero, S., Schaeffer, C., Bokhove, M., Han, L., Hebert, H., Carroni, M., Yasumasu, S., Rampoldi, L., Wu, B., Jovine, L., 2020. Cryo-EM structure of native human uromodulin, a zona pellucida module polymer. EMBO J 39, e106807. https://doi.org/10.15252/embj.2020106807

Takumida, M., 2001. Functional morphology of the crista ampullaris: with special interests in sensory hairs and cupula: a review. Biol Sci Space 15, 356–358. https://doi.org/10.2187/bss.15.356

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673–4680. https://doi.org/10.1093/nar/22.22.4673

T-Test Calculator with step by step explanation [WWW Document], n.d. URL https://www.mathportal.org/calculators/statistics-calculator/www.mathportal.org/calculators/statistics-calculator/t-test-calculator.php (accessed 2.14.23).

Vijayakumar, S., Jones, S.M., Jones, T.A., Tian, C., Johnson, K.R., 2019. Spontaneous mutations of the Zpld1 gene in mice cause semicircular canal dysfunction but do not impair gravity receptor or hearing functions. Sci Rep 9, 12430. https://doi.org/10.1038/s41598-019-48835-5

Zhao, M., Gold, L., Dorward, H., Liang, L., Hoodbhoy, T., Boja, E., Fales, H.M., Dean, J., 2003. Mutation of a conserved hydrophobic patch prevents incorporation of ZP3 into the zona pellucida surrounding mouse eggs. Mol Cell Biol 23, 8982–8991. https://doi.org/10.1128/MCB.23.24.8982-8991.2003

Published research articles related to this work

Knepper, M.I., Dernedde, J., 2022. Zona Pellucida like Domain Protein 1 (ZPLD1) Polymerization is Regulated by Two Distinguished Hydrophobic Motifs. Int. J. Mol. Sci. 2022, 23, 13894. https://doi.org/10.3390/ijms232213894

Curriculum vitae – Marie Isabell Knepper

The curriculum vitae is excluded from the online version due to data protection reasons.

Supplementary information

Full sequence of initially ordered expression vector

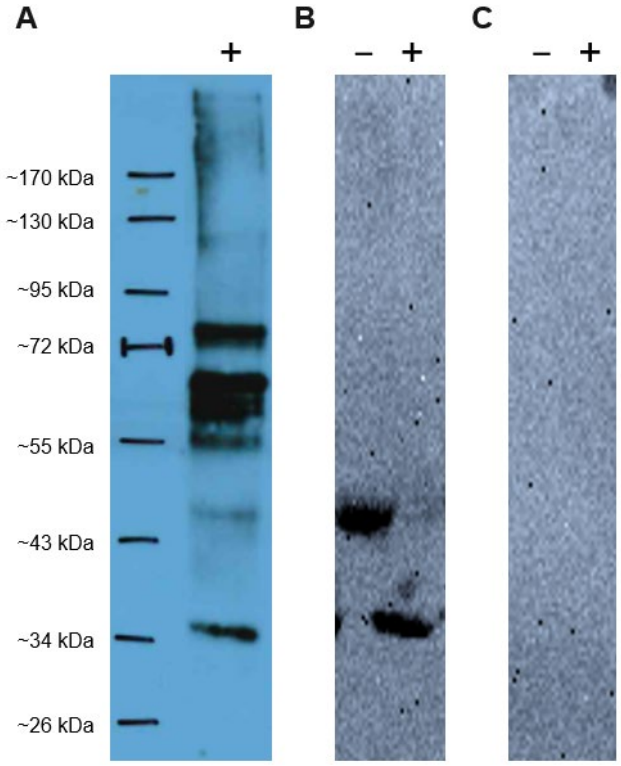
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTA
CGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACA
TCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG
CGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA
CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGA
CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGC
TAGCCACC**ATGGAACGTTTGTGTGTGATTCTCTTGCTTGTAAGTAAAACTTTCATAGCGAATGCTCAATTCAAT**
GGGTTCAACTGTGATGCCAACTTCCACAGCCGCTTCCCCGCCGAGCGGGACATCAGTGTGTATTGTGGAGTTCA
AACCATAACTCTGAAGATTAATTTTTGCCCGGTGCTTTTTTCTGGCTACACTGACACTGACCTGGCACTGAATG
GGCGTCACGGGGATGCCCACTGCAGAGGCTTCATCAACAACAACACGTTCCCCACGGTTGTGCTGTTTAGCATC
AGTCTCAACACTCTGGAGTCTTGCGGAAACTCACTGGTGGTCTCTACGGCTCATGGGCCCAACGCCTATGGGAA
TCTTTCCCTGGTCCAGATTGGGAATATAGCAGGATACATTGACACTCCGGATCCCCCGACTGTCATCAGTTACC
TACCAGGACTGCTTTACAAGTTCAGTTGTAGCTACCCACTGGAATACCTGGTCAACAACACACAGCTTGCTTCG
TCGGCCGCAGCGATATCAGTGAAGGACAGTAATGGTACTTTTGTTAGCACGCTCAATTTGCTTCTATACAACGA
CTCCTCATACATTCAGAACCTCGCCATCCCAATGTCTGGACTTACTCTGAAGACGCGAGTGTTCGCCGCTGTAA
AAGCCACCAACCTAGATAGGAGGTGGAACGTTCTGATGGACTACTGTTACACCACTCCCTCTGGAAATCCTAAT
GATGAGCTGCGATATGATCTTTTCTTTGGATGCGATAAGGACCCACAGACGACAGTGTTTGAGAACGGGAAGAG
CCAGATGGGTCGCTTCTCCTTTGAGGTGTTTCGCTTTGTCAAGCACAAAAACCAGAAAATGTCGACTGTTTTTC
TGCACTGCGTGACCAAGCTGTGTCGCTCTGATGACTGCCCCATGCTCCTGCCCATCTGCGGTAAGAGGAAGAAA
AGGGATGTGTCAGAAAGAAGCGGTGGTGCCTCAGACAATGCAGTTATAACCGCTGGACCCATCATCACCCGAAG
TGATGAAACTCCCAGTAATATATCCCAACTAGCGCAGTTAAACGGCCCTCCATTCAAAATGAATTCAGTGACGA
GCGCTTTAATTTCTGCCATCGCGATCCTCGCCGTTATGAGCATGTGTTTCTTCATCTTGTCTATGTCGCTGCTG
AGGGGAAAACAGACGCCAGCAACATCTCTATCAGGATCCCACAACCCAGCGTTCAGCCTCGAGCTCAAGCTTCG
AATTCTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGT
TCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG
GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTG
GCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACG
ACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC
AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA
GGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACA
AGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGAC
CACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTC
CGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCA
CTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGATCTATGGAGACCCCGTCCCAGCGGCGCGCCACCCGC

AGCGGGGCGCAGGCCAGCTCCACTCCGCTGTCGCCCACCCGCATCACCCGGCTGCAGGAGAAGGAGGAGGATCC
ACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCC
ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATG
GTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTG
TCCAAACTCATCAATGTATCTTAACGCGTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTT
GTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAG
ATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCG
AAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCC
GTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCG
AGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAAC
CACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA
CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTT
CAATAATATTGAAAAAGGAAGAGTCCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGA
AAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAA
GTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCC
TAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTT
ATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT
AGGCTTTTGCAAAGATCGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCA
GGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGC
CGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATG
AACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTT
GTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGC
TCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCAT
TCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGAT
CTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGAGCATGCCCGACGGCGA
GGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCA
TCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAG
CTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTT
CTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCT
GCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCG
GCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCTAGGGGGAGGCTAACTGAAACA
CGGAAGGAGACAATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGGTGTTGG
GTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGCTGGCACTCTGTCGATACCCCACCGAGACCCCATTGGG
GCCAATACGCCCGCGTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAGGCCCAGGGCTCGCAGCC
AACGTCGGGGCGGCAGGCCCTGCCATAGCCTCAGGTTACTCATATATACTTTAGATTGATTTAAAACTTCATTT
TTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGT
TCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGC
TGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCC
GAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACT
TCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGG

TTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAA
GCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACG
AGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCC
ATGCAT

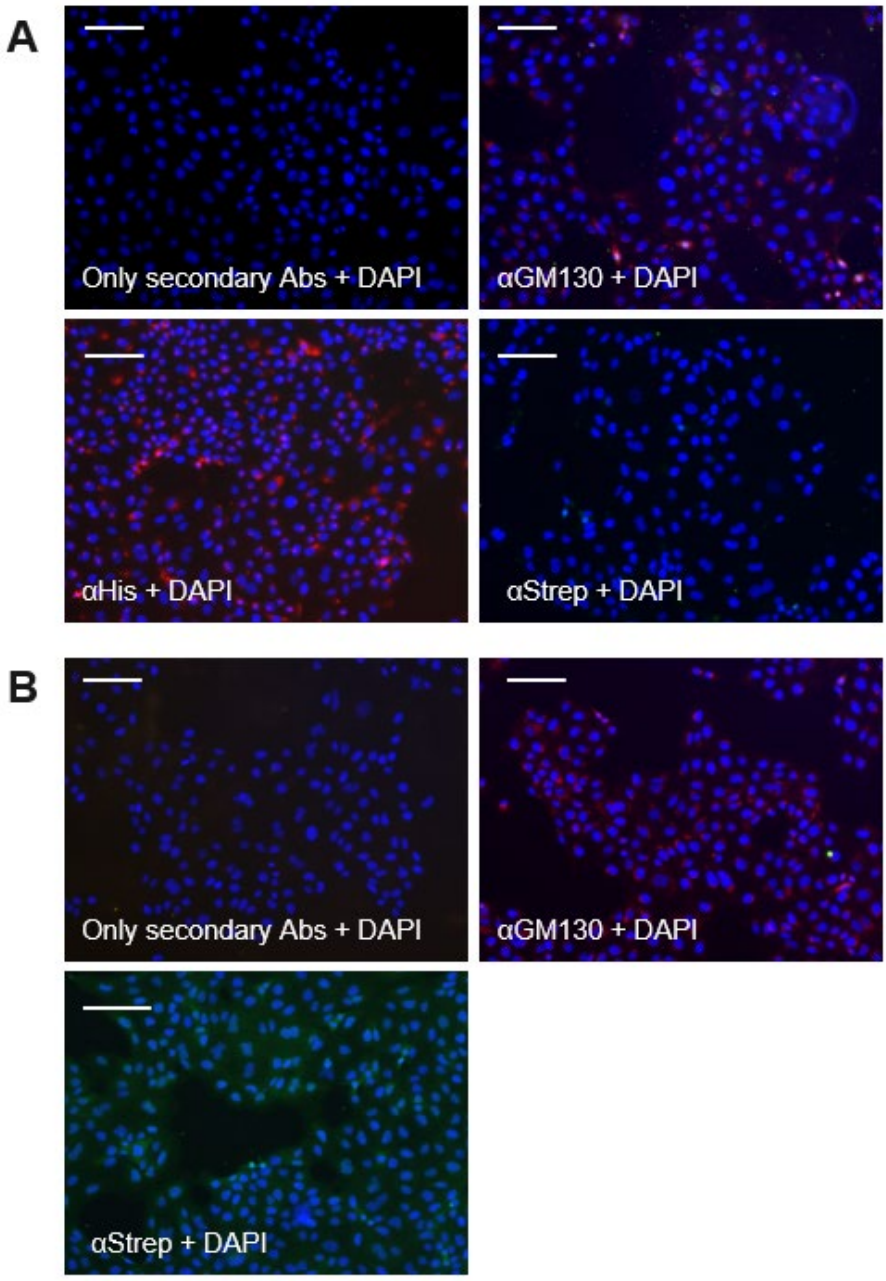
Supplementary Figure 1. Full sequence of initially ordered expression vector. The zebrafish ZPLD1 sequence (accession A0A0R4ITH5) is indicated in bold letters.

Western blot detection of PNGase F treated (+) as well as untreated (-) ZPLD1



Supplementary Figure 2. Western blot detecting PNGase F treated (+) as well as untreated (-) ZPLD1. The protein was purified from MDCK cell culture media. (A) Protein reduced and denatured in Laemmli buffer containing 2.5% mercaptoethanol at 96 °C for 5 min. Besides the 35 kDa monomer, higher molecular weight species can be observed. (B) Protein reduced and denatured in Laemmli buffer containing 5% mercaptoethanol at 96 °C for 15 min to obtain a higher portion of monomeric protein. (C) Control sample obtained from mock transfected MDCK cells.

Control immunofluorescence images of untransfected MDCK cells



Supplementary Figure 3. Control immunofluorescence images of untransfected MDCK cells. Cells treated with: a mixture of all used secondary antibodies, antibodies against the Golgi marker GM130, and tag antibodies (His/Strep) (see Materials and Methods). Alle samples were additionally stained with DAPI. (A) Cells were not permeabilized; (B) Cells permeabilized using Triton X-100. Bar, 50 µm.

Raw data of ZPLD1 average polymer surface quantification

Supplementary Table 1: Raw Data for average polymer surface quantification of presented PIC experiment.

Picture	Count	Total Area	Average size	%Area
DMSO_x10_I	17	23181	1363,588	1,606
DMSO_x10_II	20	31174	1558,7	2,16
DMSO_x10_III_c2-1.JPG	22	15698	1432,438	1,088
DMSO_x10_IV	15	18091	1206,067	1,253
DMSO_x10_V_c2-1.JPG	22	31369	3869,361	2,173
DMSO_x10_VI	22	21871	994,136	1,515
DMSO_x10_VII	13	19706	2704,433	1,365
DMSO_x10_VIII	17	20531	1207,706	1,422
DMSO_x10_IX	14	19169	1369,214	1,328
PIC0.5ul_x10_I	22	18855	857,045	1,306
PIC0.5ul_x10_II	16	22760	1422,5	1,577
PIC0.5ul_x10_III	10	11702	1170,2	0,811
PIC0.5ul_x10_IV	19	14257	750,368	0,988
PIC0.5ul_x10_V	11	10721	974,636	0,743
PIC0.5ul_x10_VI	13	14168	1089,846	0,981
PIC0.5ul_x10_VII	16	12556	784,75	0,87
PIC0.5ul_x10_VIII	17	16095	946,765	1,115
PIC0.5ul_x10_IX	16	14532	908,25	1,007
PIC0.75ul_x10_I	13	12457	958,231	0,863
PIC0.75ul_x10_II	16	14191	886,938	0,983
PIC0.75ul_x10_III	24	12716	529,833	0,881
PIC0.75ul_x10_IV	13	10186	783,538	0,706
PIC0.75ul_x10_V	12	13761	2485,445	0,953
PIC0.75ul_x10_VI	28	13784	492,286	0,955
PIC0.75ul_x10_VII	11	6122	556,545	0,424
PIC0.75ul_x10_VIII	12	6787	565,583	0,47
PIC0.75ul_x10_IX	17	11006	647,412	0,762
PIC1ul_x10_I	4	2287	571,75	0,158
PIC1ul_x10_II	3	2031	677	0,141
PIC1ul_x10_III	11	5688	517,091	0,394
PIC1ul_x10_IV	13	7173	551,769	0,497
PIC1ul_x10_V	10	4925	492,5	0,341
PIC1ul_x10_VI	10	5433	543,3	0,376
PIC1ul_x10_VII	15	8149	543,267	0,565
PIC1ul_x10_VII	13	5382	414	0,373
PIC1ul_x10_IX	20	12154	607,7	0,842

Supplementary Table 2: Raw Data for average polymer surface quantification of presented single proteases experiment.

Picture	Count	Total Area	Average size	%Area
AEBSF_x10_I	9817	892,46	0,68	32,902
AEBSF_x10_II	7889	1127,00	0,547	13,363
AEBSF_x10_III	9414	1046,00	0,652	60,154
AEBSF_x10_IV	8830	1471,67	0,612	55,965
AEBSF_x10_V	10510	1167,78	0,728	48,067
AEBSF_x10_VI	6811	756,78	0,472	35,932
AEBSF_x10_VII	9453	675,21	0,655	31,333
AEBSF_x10_VIII	5520	1104,00	0,382	40,222
AEBSF_x10_IX	11117	1588,14	0,77	93,096
B_x10_I	21007	1105,63	1,455	57,749
B_x10_II	32807	1491,23	2,273	22,679
B_x10_III	28891	1605,06	2,001	29,236
B_x10_IV	42050	1828,26	2,913	32,108
B_x10_V	33033	1651,65	2,288	28,623
B_x10_VI	27334	976,21	1,894	25,889
B_x10_VII	20140	1765,84	1,396	89,895
B_x10_VIII	38622	1287,40	2,676	11,018
B_x10_IX	28758	1369,43	1,992	31,373
E64_x10_I	34101	1623,86	2,362	30,283
E64_x10_II	17541	1349,31	1,215	23,951
E64_x10_III	39887	1661,96	2,763	38,076
E64_x10_IV	23527	1568,47	1,63	26,936
E64_x10_V	26464	1470,22	1,833	40,617
E64_x10_VI	26094	2174,50	1,808	29,369
E64_x10_VII	29475	1052,68	2,042	26,261
E64_x10_VIII	30163	1160,12	2,09	13,84
E64_x10_IX	30513	1271,38	2,114	31,57
P_x10_I.JPG	39191	1351,41	2,715	46,99
P_x10_II	40248	1387,86	2,788	83,054
P_x10_III	34476	2154,75	2,388	146,441
P_x10_IV	27778	1462,00	1,924	32,998
P_x10_V	40873	1857,86	2,831	119,322
P_x10_VI	24598	1892,15	1,704	70,305
P_x10_VII	23788	1132,76	1,648	66,392
P_x10_VIII	13763	983,07	0,953	101,002
P_x10_IX	33950	1997,06	2,352	121,817
untreated_x10_I	26722	1272,48	1,851	108,699
untreated_x10_II	26946	1796,40	1,867	60,736
untreated_x10_III	23071	1153,55	1,598	51,398
untreated_x10_IV	27380	1521,11	1,897	104,16
untreated_x10_V	24653	1760,93	1,708	106,073

untreated_x10_VI	35447	1417,88	2,456	53,37
untreated_x10_VII	33572	2098,25	2,326	97,789
untreated_x10_VIII	49399	1764,25	3,422	70,373
untreated_x10_IX	30237	1314,65	2,095	165,426
DMSO_x10_I	37422	1496,88	2,592	21,268
DMSO_x10_II	34247	1556,68	2,372	35,834
DMSO_x10_III	21066	1170,33	1,459	18,805
DMSO_x10_IV	28975	2228,85	2,007	18,185
DMSO_x10_V	41763	2198,05	2,893	44,615
DMSO_x10_VI	38248	1366,00	2,65	20,001
DMSO_x10_VII	26709	1214,05	1,85	24,232
DMSO_x10_VIII	31516	1575,80	2,183	32,095
DMSO_x10_IX	22542	1734,00	1,562	43,099