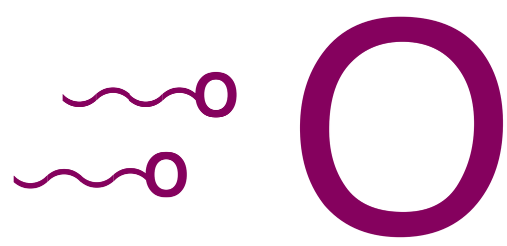


Crystal structure of the most conserved domain of mouse sperm receptor ZP3

Magnus Monné*, Ling Han*, Thomas Schwend, Sofia Burendahl and Luca Jovine#

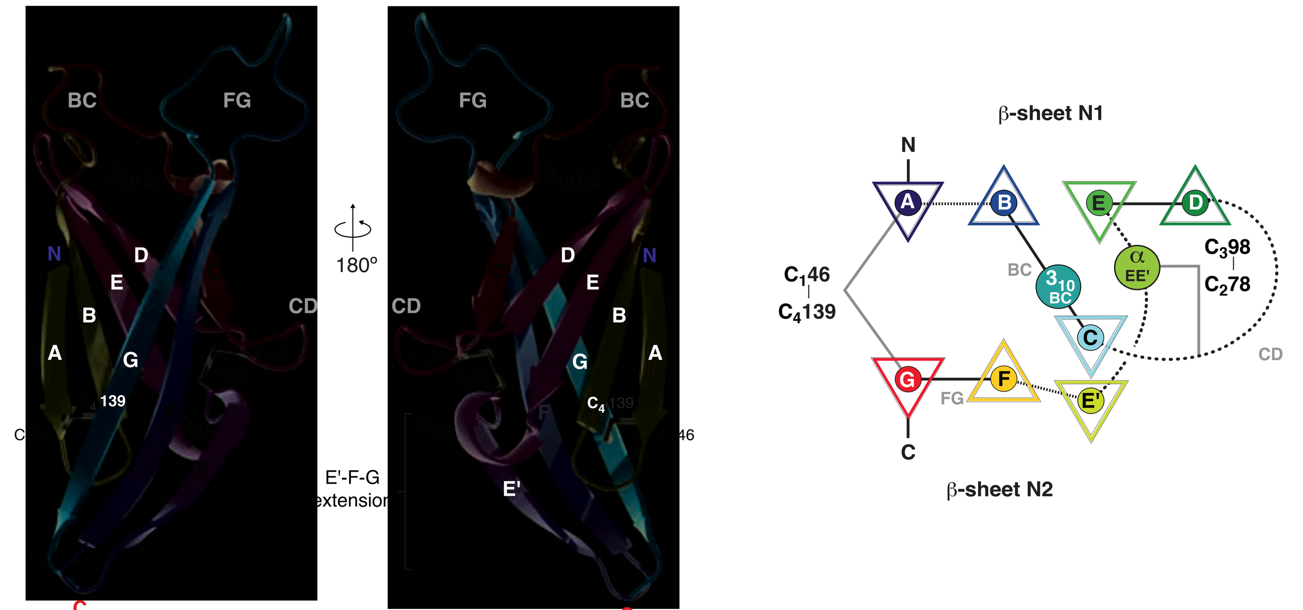
Department of Biosciences and Nutrition, Karolinska Institutet, Hälsovägen 7, 141 57 Huddinge, Sweden.

*Have contributed equally. #Corresponding author: luca.jovine@ki.se.

Abstract

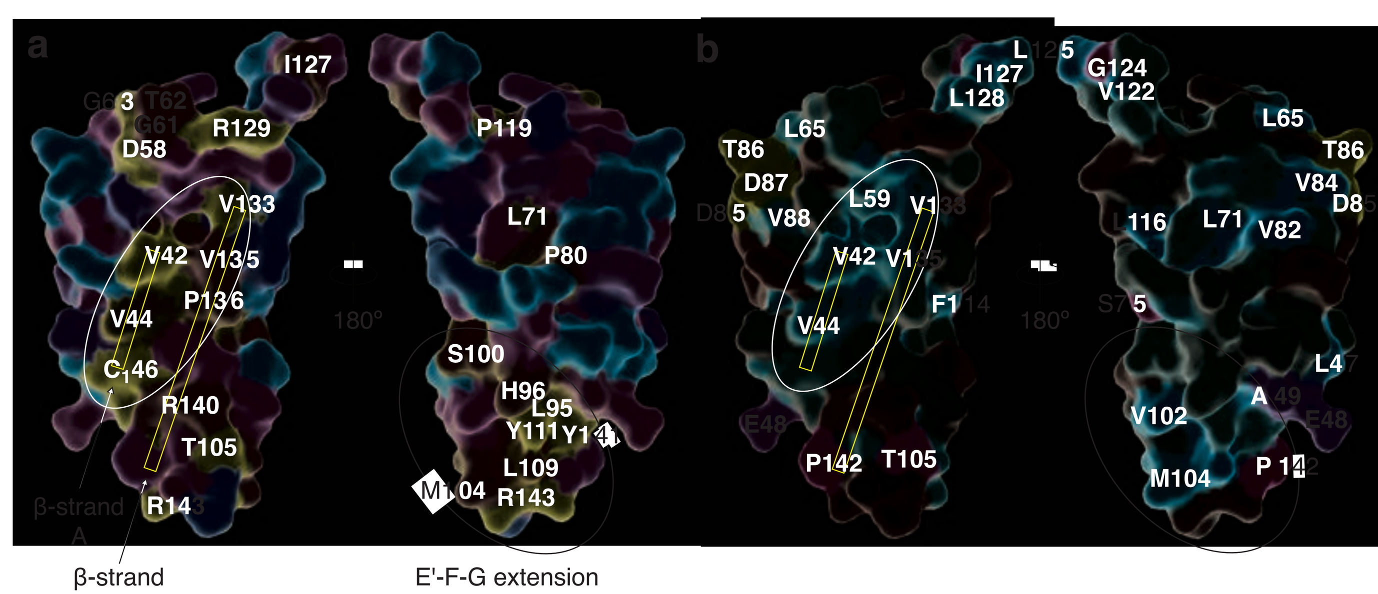
The first step of mammalian fertilization is species-specific recognition between the egg extracellular matrix (zona pellucida, ZP) and sperm. The ZP is built up of filaments that are composed of sperm receptors ZP2 and ZP3 and are cross-linked by ZP1. The first two subunits are crucial for reproduction since mice lacking either of their genes produce oocytes without a ZP and are completely infertile. All ZP components contain a polymerization 'ZP domain' module that can also be found in egg coats of non-mammalian species as well as many other secreted eukaryotic proteins. Here we describe the 2.3 Å resolution structure of the conserved N-terminal part (ZP-N) of mouse primary sperm receptor ZP3, which was previously suggested to constitute a domain of its own. This is the first atomic structure of a ZP domain protein and of a conserved vertebrate protein that is essential for fertilization and directly involved in egg-sperm binding. The ZP-N fold defines a new immunoglobulin superfamily subtype with a β-sheet extension characterized by an E'-strand and an invariant tyrosine residue implicated in polymerization. The structure reveals the presence of ZP-N repeats within the N-terminal region of ZP2 and other egg coat proteins, with implications for overall ZP architecture, the post-fertilization block to polyspermy and speciation. The molecular basis of a deafness-causing mutation within the ZP-N domain of human inner ear protein α-tectorin is examined in detail, showing that the structure provides an important framework to understand human diseases caused by mutations in ZP domain proteins. Moreover, the ZP-N domain fold could serve as template for structure-based design of targeted non-hormonal contraceptives.

Overall structure and topology

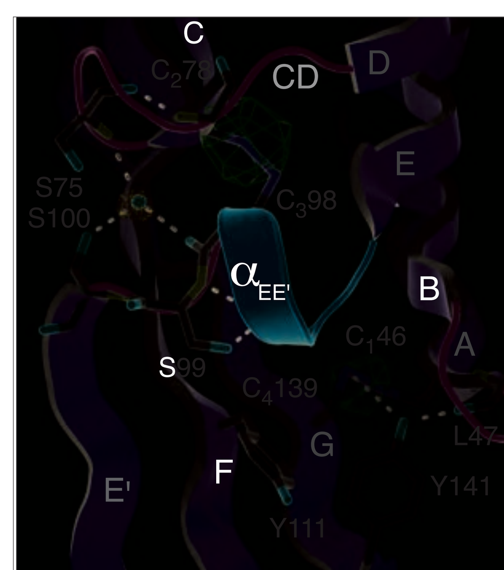


ZP domain proteins have been found to be notoriously difficult to over-express since they polymerize and contain disulfide bonds, N- and O-linked glycosylations. After extensive trials we managed to express the N-terminal (ZP-N) domain of mouse ZP3 as a fusion protein with Maltose Binding Protein (MBP) in an *Escherichia coli* strain able to form disulfide bonds. The fusion protein could be purified, crystallized and diffraction data was collected. The structure could be solved with molecular replacement with MBP and the model was refined to 2.3 Å resolution. Unexpectedly, the about 100 amino acid residue ZP-N domain is an immunoglobulin-like β-sandwich. The hydrophobic core is clamped together by the two canonical disulfide bonds of ZP-N domains with the connectivity C_1-C_4 and C_2-C_3. The C-strand is continued by the E'-strand, an arrangement not seen before, that is a part of the E'-F-G β-sheet that extends below the hydrophobic core. Therefore ZP-N domain belongs to a new immunoglobulin superfamily subtype.

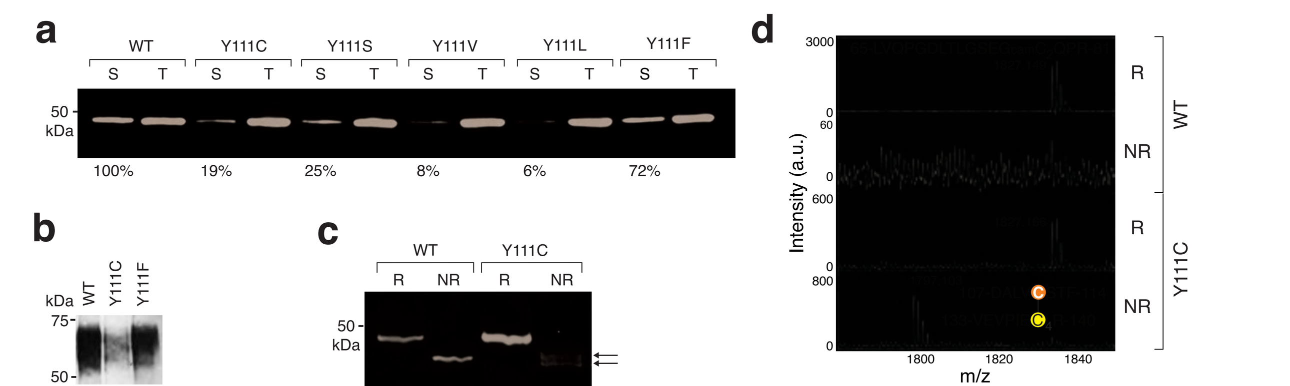
ZP-N domain interaction sites



Surface representation with conservation increasing from red to blue (**a**) and hydrophobicity increasing from white to red with N-glycosylation, protease cleavage and RGD sites in green, yellow and blue, respectively (**b**), indicate that the MBP facing side (white oval) and the E'-F-G extension (black oval) are conserved, hydrophobic and not necessarily exposed in filaments which makes them potential sites involved in polymerization.

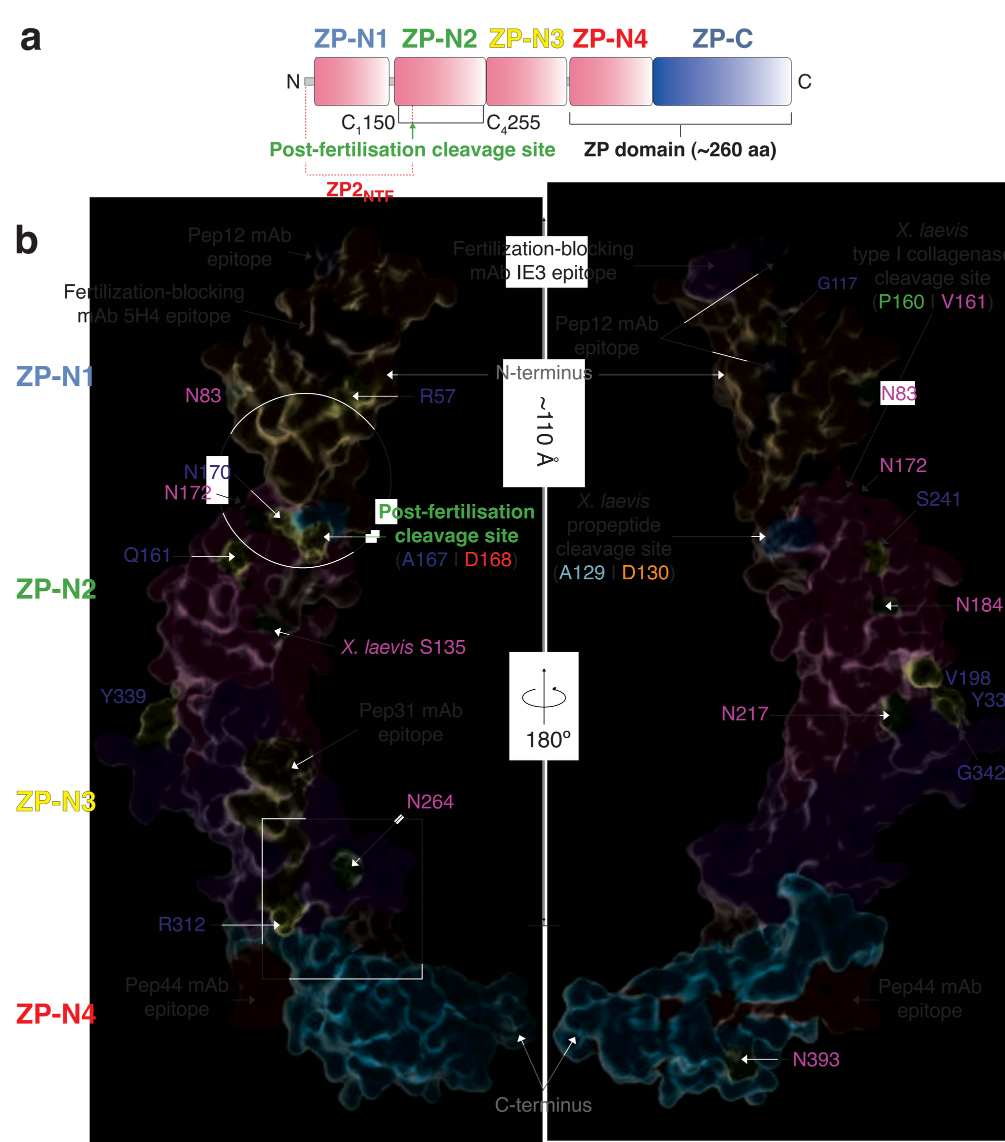


The E'-F-G extension contains two tyrosines that, apart from the cysteines, are nearly invariant. When the homologues ZP-domain in α-tectorin has the mutation Y1870C the function for the tectorial membrane in the inner ear is disrupted and causes deafness. The corresponding tyrosine in the ZP-N structure of mouse ZP3 is Y111, situated in proximity to $C_4$139 indicating that such a mutation could interfere with the native disulfide bond formation.

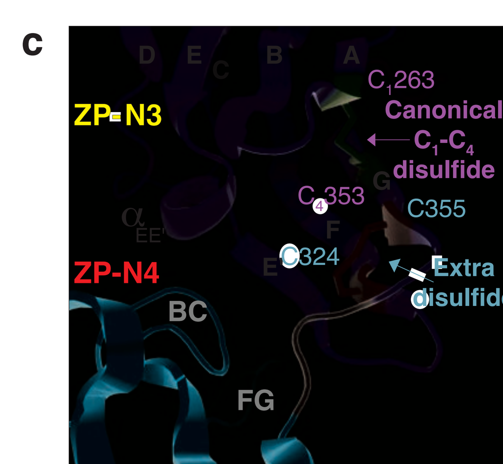


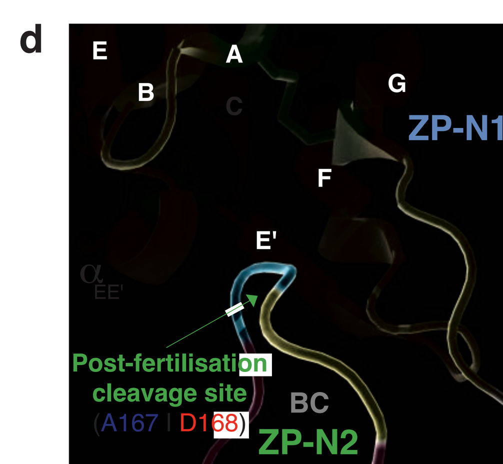
Mutations of Y111. Analysis of soluble (S) and total (T) protein of Y111 mutants shows reduced solubility when expressed in *E. coli* (**a**) and reduced secretion when expressed in CHO-K1 cells (**b**). However, such a decrease leading to haploinsufficiency cannot be the explanation for the disease since the tectorial membrane is not disrupted in animals where the α-tectorin gene is deleted on only one allele. Immunoblot analysis of the Y111C mutation in reducing (R) and non-reducing (NR) conditions shows that the mutant protein exist in two forms (**c**), of which one contain the disulfide C111-$C_4$139 (**d**). Thus the small fraction of mutant protein that gets secreted acts as a dominant negative by incorporating into the filaments and preventing their elongation. This proves the involvement of the E'-F-G extension in polymerization.

ZP-N domain repeats

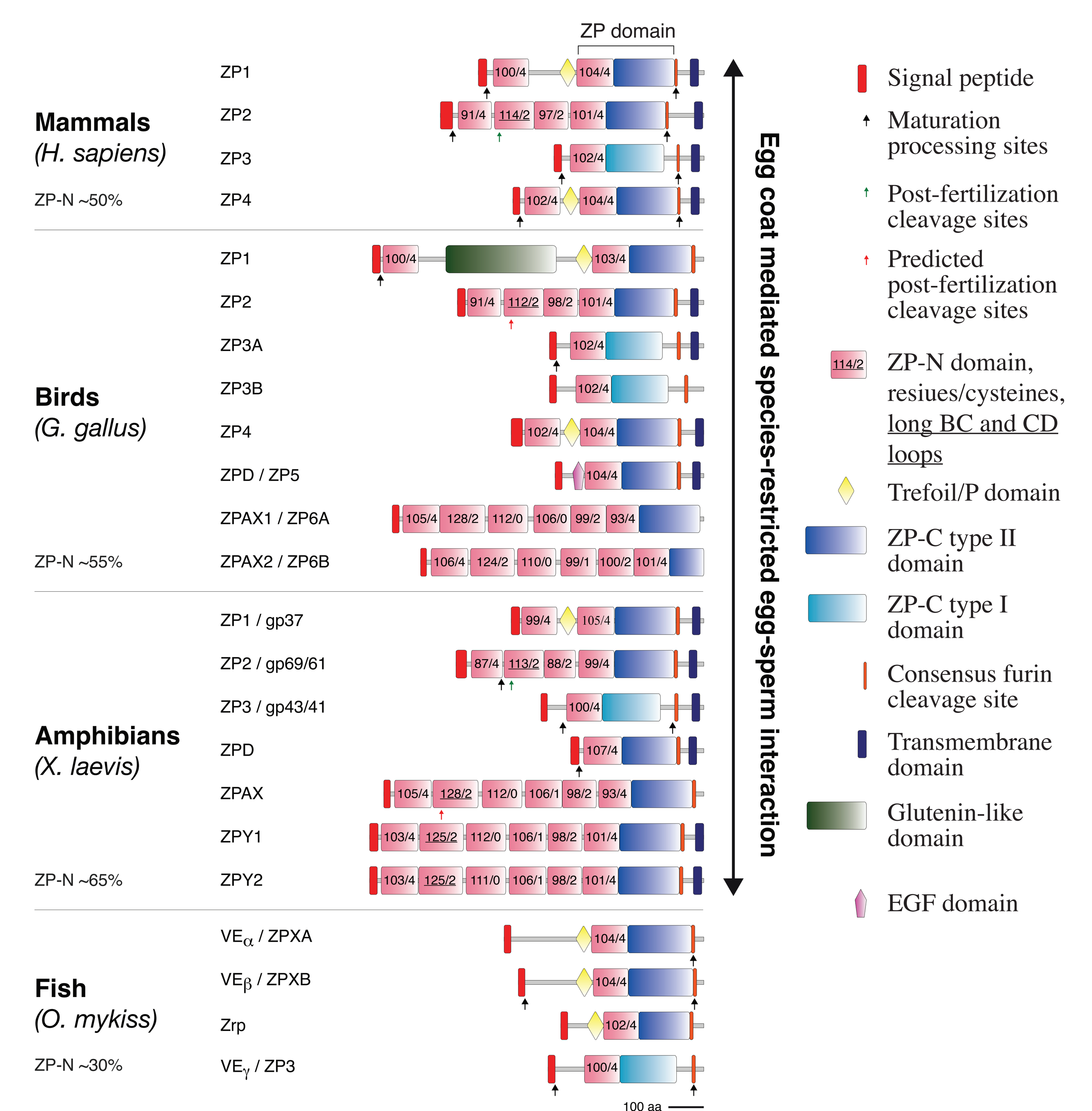


In mouse secondary sperm receptor ZP2 we identified 3 additional ZP-N domains N-terminal of the C-terminal ZP-domain (ZP-N4 and ZP-C) (**a**) by using a structural based approach. This idea has earlier been hypothesized based on sequence but the homology modelling of the 4 ZP-N repeat of ZP2, coloured from blue to red (**b**), concludes it and is structurally sound, as exemplified by the formation of a potential extra, non-canonical disulfide bond (**c**). In addition, sites known to be exposed are mapped onto the model - glycosylation and positively selected sites in violet and dark blue, respectively, and eptiopes for monoclonal antibodies and protease cleavage sites. Interestingly, the ZP-N1-N4 assumes an extended rod-shaped structure, spanning about 110 Å that matches the size of protrusions from isolated ZP filaments imaged by electron microscopy.





The post-fertilization block to polyspermy is achieved by site-specific proteolysis between the amino acids corresponding to A167 and D168 in mouse ZP2. This site is found in the BC loop of ZP-N2 (**d**), which upon cleavage could alter the orientation of ZP-N1 relative to ZP-N2 and lead to a formation of an "adhesive zipper" that would trigger contacts between the ZP-filaments, leading to the observed compaction of the ZP and form a physical barrier to sperm penetration.



Extension of the structure-based analysis reveals that the ZP-N fold accounts for >50% of the structure of all egg coats from frog to human and no additional ZP-N repeats are found in any other ZP domain protein not involved in fertilization. This raises an unexpected parallel with invertebrate gamete recognition proteins, where rapid evolution of repeated domains is implicated in speciation.

This work has been published in *Nature* **456**:653-657 (2008). For further information see http://www.jovinelab.org.