First publ. in: The FASEB Journal 20 (2006), 2, pp. 317-319

Disparate evolution of prion protein domains and the distinct origin of Doppel- and prion-related loci revealed by fish-to-mammal comparisons

Eric Rivera-Milla,*[,1] Birgit Oidtmann,[†,1,2] Cynthia H. Panagiotidis,[‡] Michael Baier,[§] Theodoros Sklaviadis,[‡] Rudolf Hoffmann,[†] Yi Zhou,[‖] Gonzalo P. Solis,* Claudia A. O. Stuermer,* and Edward Málaga-Trillo*[,3]

*Department of Biology, University of Konstanz, Konstanz, Germany; [†]Institute of Zoology, Fish Biology and Fish Diseases, University of Munich, Munich, Germany; [‡]School of Pharmaceutical Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece and INA-CERTH, Thermi, Thessaloniki, Greece; [§]Project Neurodegenerative Diseases, Robert Koch Institute, Berlin, Germany; and [‖]Division of Hematology/Oncology, Children's Hospital Boston and Dana-Farber Cancer Institute, Boston, Massachusetts, USA

SPECIFIC AIMS

Our current understanding of prion biology and disease is largely based on studies performed on mammals, yet basic questions about the physiological function of prion proteins (PrPs) and the molecular nature of prion disorders remain elusive. To facilitate the establishment of non-mammalian models for prion research, we characterized the sequence, structural, and genomic homology between fish and other vertebrate PrPs, and analyzed the distinct molecular mechanisms that shaped the evolution of vertebrate PrP domains.

PRINCIPAL FINDINGS

1. Teleost fish possess duplicated PrPs, which are the genetic, structural, and syntenic homologues of mammalian PrP and doppel

We provide new sequence data documenting the existence of two orthologous *PrP* loci in bony fish (*PrP-1* and *-2*), which display extensive variation in their length and amino acid composition, and which are highly expressed in adult and developing fish brains. Despite the low sequence similarity with their mammalian counterparts, fish PrPs share all the expected PrP structural landmarks, such as an N-terminal repetitive region, a highly conserved hydrophobic domain, and a predicted C-terminal globular domain containing two β-strands, three α-helices, and a GPI-anchoring motif (**Fig. 1*A***). Our genomic analysis in zebrafish and *Fugu* shows that *PrP-1* and *-2* map to different chromosomes which are mosaically syntenic to mammalian *PrP* chromosomes. Directly adjacent to each *PrP* there is a *PrP*-related locus (*PrP-rel-1* and *-2*) encoding a short GPI-anchored polypeptide with the unique PrP hydrophobic domain (Fig. 1*A*). Cladistic analysis supports the idea that *PrP* and *PrP-rels* arose as tandem duplicates—a situation reminiscent of that of mammalian *Prnp* and *Doppel* (*Dpl*)— and later duplicated in block as the consequence of a large chromosomal/genome duplication in a teleost fish ancestor (**Fig. 2**). Due to the subsequent differential loss of paralogues in the fish genomes, the mosaic synteny is only detectable when both duplicated fish regions are considered together.

2. Vertebrate PrP domains have evolved as modules following independent evolutionary dynamics

Our structural analyses across all vertebrate classes reveals that the N- and C-terminal protein moieties have evolved independently from one another following their own rules and different evolutionary pressures: while the former underwent differential (class-specific) expansion-degeneration cycles in its repetitive domains, the latter retained its basic globular structure despite high sequence divergence.

---

[1] These authors contributed equally to this work.

[2] Current address: Centre for Environment, Fisheries and Aquaculture Science, Weymouth Laboratory, Barrack Rd., The Nothe, Barrack Rd., Weymouth, Dorset DT48UB, UK.

[3] Correspondence: Department of Biology, University of Konstanz, Konstanz 78457, Germany. E-mail: edward.malaga@uni-konstanz.de

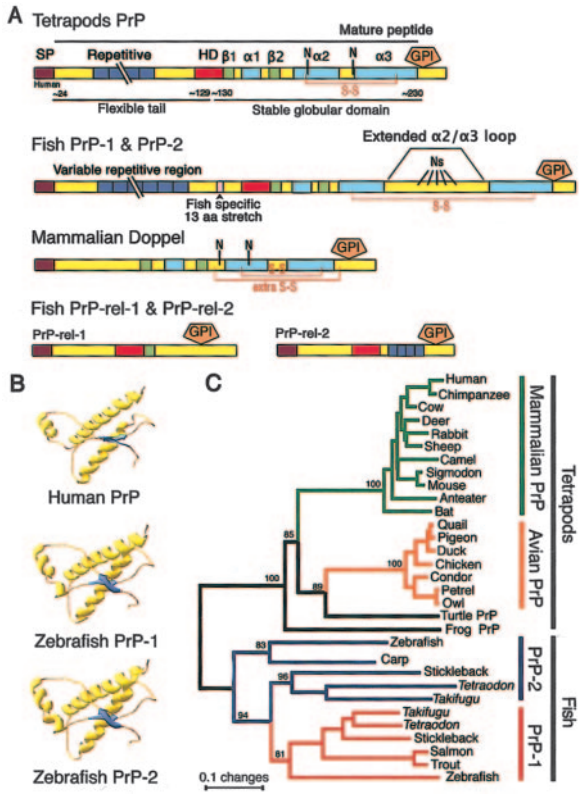

Figure 1. PrP diversity in vertebrates. *A)* Landmark proteins motifs of PrP, Doppel and PrP-rel in tetrapods and bony fish. Repeats are shown in blue, hydrophobic domains (HD) in red, β-strands in green, α-helices in cyan and the fish specific 13 aa motif in pink. SP, signal peptide; N, glycosylation sites. Breakpoints at repetitive regions indicate length variation. *B)* 3D structures of human and zebrafish (predicted) PrP globular domains. *C)* Evolutionary relationships among vertebrate PrP globular domains using distance and Neighbor-Joining methods (bootstrap values are shown at relevant nodes).

3. Structural conservation of the globular domain from fish to mammals suggests that it harbors an essential biological activity

Despite having low sequence similarity scores, fish and human PrPs share the same structural landmarks and secondary structure predictions, yielding very similar 3D-models (Fig. 1*B*). Our analysis of amino acid replacements between zebrafish and human indicates that almost all changes occurring within α-helices and β-strands are structurally equivalent, causing no major effect on the overall predicted folding of the polypeptide. The extensive sequence divergence between vertebrate classes contrasts with the strong conservation observed within a single class. The branching pattern in our phylogenetic analyses (Fig. 1*C*) suggests that a shift in selective pressure generated two distinct phases of molecular evolution at the globular domain: an early divergent phase where weak negative selection and rapid substitution shaped the typical proteins of each vertebrate class, and a later stasis phase where strong negative selection favored intraclass conservation. While our analysis predicts a conserved function for all vertebrate globular domain, it is not yet possible to ascertain whether misfolding is also a widely conserved property of all PrPs or whether it is restricted to mammals.

4. Differential expansion/degeneration of repetitive elements at the N-terminal region

Detailed analysis of the intricate fish PrP repetitive domains led us to uncover a complex pattern consisting of two basic core units (here called type A and B repeats), which are indeed present (although not equally represented) in all vertebrate classes. Thus, the described previously copper binding mammalian octarepeats belong to the B type, and they are preceded by the newly identified type A repeats (**Fig. 3*A***). Analysis of PrP repeat variation across vertebrate classes shows that A and B repeats underwent different events of expansion and divergence specific to each class, resulting in the preferential expansion of either one of the two types of repeats (Fig. 2*B*). Unlike all other vertebrate classes, mammals are the only organisms that show expansion of B repeats, (Fig. 3*C*). The large degree of variability between the repetitive regions across vertebrate classes suggests that these elements might serve a secondary, perhaps structural role subordinated to the biological function of the globular domain.

CONCLUSIONS AND SIGNIFICANCE

Because several fish PrP-like molecules had been reported by us and others, it was necessary to unambiguously assess their relationships. Our data clearly establish the correspondence between mammalian and fish PrPs using structural, phylogenetic, and syntenic analyses. While PrPs are often regarded as highly conserved proteins, our data emphasizes the fact that this observation does not apply beyond mammals and that there is a considerable amount of PrP molecular variation that needs to be considered and analyzed. Moreover, we find that vertebrate PrPs have evolved as two separate modules, with the repetitive region and the globular domain following independent patterns of evolution, which strongly implies at least two very different functional properties of the native molecule. The idiosyncratic evolution of different repetitive regions in every vertebrate class underscores a very flexible role for this domain, making it necessary to reexamine previous assumptions about conserved roles in, for example, copper metabolism, or cellular processing. More generally, as the physiological role of PrP remains largely unknown, our identification of evolutionarily conserved structural features across vertebrate PrPs provides a starting point toward experimentally dissecting the relative contributions of these protein motifs to the biological function of PrPs. Moreover, parallel characterization of fish PrPs by us and others will bring us closer to establishing the zebrafish as an animal model for the study of TSE and other neurodegenera-

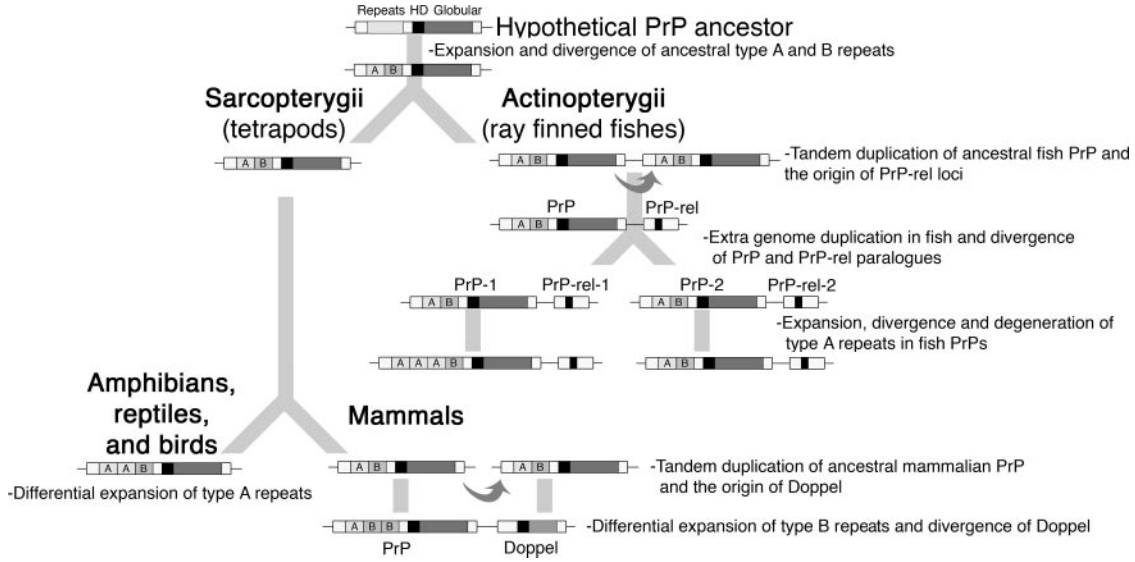

Figure 2. Proposed major events along vertebrate PrP evolution. Fish and mammalian PrPs underwent the independent expansion of A and B repeats, respectively. Upon their independent origin by gene duplication, fish PrP-rel and mammalian doppel differentially lost PrP structural features. Repetitive regions are shown in light gray, hydrophobic domains (HD) in black, and globular domains in dark gray.

tive diseases. Likewise, identification of PrP homologues in more basal vertebrates such as sharks and jawless fish, or the cephalochordate *Amphioxus*, might help us uncover the origin of the PrP fold and the rules of misfolding and its pathogenic properties, so far only studied in mammals and yeast. FJ

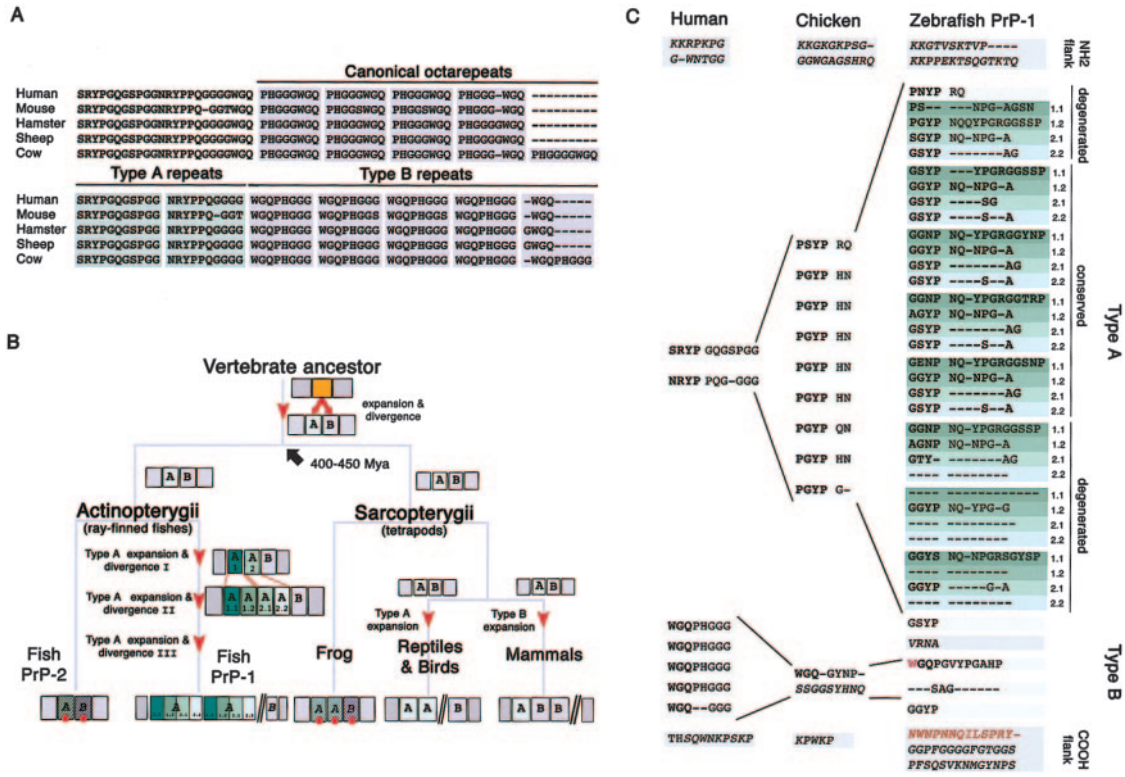

Figure 3. Idiosyncratic evolution of PrP repetitive domains. *A)* Relative occurrence of A and B repeats between vertebrate classes compared to the canonical mammalian octarepeats. *B)* Hypothetical scenario for the differential expansion of vertebrate A and B repeats from an ancestral proto-repeat (orange box). Repeat boxes with a red dot indicate degenerated repeats in fish PrP-2 and frog PrP. *C)* Vertical alignment showing the structural correspondence between A and B repeats of human, chick and zebrafish PrP-1s. Gaps are denoted by dashes. A tryptophan (W) residue possibly lost in a zebrafish B repeat (PrP-1) is shown in magenta. The fish 13 aa specific motif is highlighted in red.