Zwilling-A and -B, Two Related Myelin Proteins of Teleosts, Which Originate from a Single Bicistronic Transcript

Karin Schaefer and Christian Brösamle

Department of Biochemistry, Ludwig-Maximilians Universität, München, Germany

Myelination, the ensheathment of axons by membranes of highly specialized glial cells, has been a crucial innovation during early vertebrate evolution. It enables high nerve signal conduction velocities, while maintaining nervous system size and energy requirements at moderate levels. Consequently, myelination has been conserved in all extant gnathostome vertebrates. In a genomewide mRNA expression screen, we identified several novel neural crest and myelin-specific transcripts in the zebrafish (Danio rerio). Here, we describe the characterization of two proteins, Zwilling-A and -B (ZwiA and ZwiB), which are exclusively expressed in myelinating glia of teleosts. They are structurally homologous and are translated from a common, bicistronic transcript. No similarities to sequences or domains of other proteins were detected. Analysis of phylogeny, genomic organization, and genomic synteny suggests that the zwi gene has appeared soon after the teleost-specific genome duplication event and evolved under conservative selective pressure. We hypothesize that ZwiA and ZwiB serve important physiological functions in teleost myelin.

In a molecular screen comparing global gene expression patterns between wildtype zebrafish larvae and myelination-deficient homozygous mutant soxl10/colourless\(^2\) larvae by microarray hybridization, we identified two previously unknown myelin-specific transcripts (L. Orosco, K. Schaefer, M. E. Halpern, C. Brösamle; unpublished data). One of these transcripts was represented by EST AW281753, the 3'-end sequence of a cDNA library constructed from adult brain mRNA. In order to derive a conceptual translation of the corresponding protein, we reconstructed a full-length cDNA by clustering other available expressed sequence tag (EST) sequences with homology to the cDNA and support axons (Simons and Trajkovic 2006).

For permissions, please e-mail: journals.permissions@oxfordjournals.org

Key words: myelin, evolution, zebrafish, genome duplication.

E-mail: christian.broesamle@med.uni-muenchen.de.

doi:10.1093/molbev/msn298
Advance Access publication January 6, 2009

© The Author 2009. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org
pos. 19 and 50 in ZwiB), and cAMP- and cGMP-dependent protein kinase (CAMP; threonine pos. 81) are present in Zwilling proteins (fig. 1B). In silico prediction of secondary structures suggested stretches of beta-sheet conformations interrupted by random coils between residues 10 and 70 for both ZwiA and ZwiB. The amino terminus appears inherently unstructured, whereas the last approximately 12 amino acids exhibit a certain propensity to adopt an alpha-helical structure. Overall, the results of the various prediction algorithms used, showed considerably variability, indicating that Zwilling proteins may not exist in strictly defined secondary structures but rather as intrinsically unstructured protein chains that can present in multiple conformations. Zwilling mRNA in situ hybridization shows the typical pattern of transcripts expressed in myelinating glia cells. Expression starts during late embryonic stages between 2 and 3 days postfertilization (dpf) in a few differentiating oligodendrocytes in the hindbrain (fig. 1D). Simultaneously, the first Schwann cells along the posterior lateral line nerve are positive for zwilling mRNA (inset). At 4 dpf, the full expression pattern of a typical myelin transcript (Brösmäle and Halpern 2002) has been reached: Many oligodendrocytes are visible in the hindbrain and ventral spinal cord, and Schwann cells of several branches of the peripheral nervous system, most notably the lateral line system and cranial nerves, can be identified (fig. 1E–F). After 4 dpf, this pattern essentially stays stable, only additional myelinating cells are added to the respective structures (data not shown).

Database searches using ZwiA and ZwiB protein sequences and DNA sequences of the corresponding ORFs yielded numerous hits in other species that turned out to be all of the teleost clade. Sequence alignment showed amino acid identities between 93% and 62% (fig. 2A). Conservation was particularly high at the amino- and carboxy-termini, mostly reflecting selective pressure to maintain the myristoyl and CAMP phosphorylation functions. Phylogenetic analysis of the protein alignment resulted in a tree that accurately reflects the known phylogenetic relationships (fig. 2B). For all five species for which significant genomic data are available (Danio rerio, Gasterosteus aculeatus, Oryzias latipes, Fugu rubripes, and Tetraodon nigroviridis), two homologous sequences could be identified that segregated clearly within their appropriate branch of the phylogenetic tree. For species without genomic data (Gillichthys mirabilis, Pimephales promelas, Oncorhynchus mykiss, Salmo salar, and Astatotilapia burtoni) either one or two homologues were found, which again clearly segregated with either the ZwiA or ZwiB branch (fig. 2B). Calculation of average $K_s$/$K_a$ values that rate nonsynonymous versus synonymous nucleotide substitution rates ($K_s$/aZwiA = 0.062 ± 0.037 SD; $K_s$/aZwiB = 0.069 ± 0.061) reveals significant conservative evolutionary pressure on both Zwilling proteins. Genomic analysis of the $zwi$ loci revealed high conservation, also on the level of gene structure. Percent–identity plots illustrate the high conservation of both ORFs. Additional sequence conservation exists in the 5’ untranslated region, the
untranslated region between the zwiA and zwiB ORFs, and in a small region of the intron (fig. 2C). Genomic conservation extends beyond the zwi locus itself. All zwi genes are nested within the first intron of one of the orthologues of mammalian RUNDC3A, which itself resides in a cluster of conserved synteny (fig. 2D). Examination of the corresponding zebrafish paralogues shows that none of them has any recognizable gene residing in its first intron. Moreover, their first introns are always significantly shorter than the RUNDC3A orthologues that contain the zwi locus. In all examined species, a conserved syntenic relationship exists over at least four loci 5' to zwi. In perciform fishes, this is also true for the 3' direction. In the cyprinid Danio rerio, the 5' end of the RUNDC3A locus represents the end of the syntenic cluster.

With two peptides translated from a single mRNA, the zwilling gene possesses a rather unusual structure. Moreover, its two ORFs are highly similar and code for proteins that likely have similar functions. This suggests a tandem duplication event of a small DNA segment, resulting in a bicistronic transcript. Unlike in other genes of a similar arrangement, such as Drosophila adh/adhr (Broga and Ashburner 1997) or zebrafish wnt8 (Lekven et al. 2001) where nonduplicated, monocistronic forms are known in other species, our phylogenetic analysis suggests that zwilling appeared as an insertion already in its duplicated form. The presence of zwilling in species of the Ostariophysi (D. rerio, P. promelas), Proacanthopterygii (G. aculeat), and Acanthopterygii (F. rubripes, O. mykiss, and others) clades, and its genomic location in the first intron of only one of the teleost orthologues of RUNDC3A, suggests an insertional event early in the teleost lineage, but after the genome duplication at the base of the teleost radiation (Amores et al. 1998; Prince et al. 1998; Taylor et al. 2001; Hurley et al. 2007). The origin of this ancestral zwilling gene remains elusive.

The zwilling gene, the structure of its mRNA, and the sequences of its protein products have been conserved over
more than 200 My. \( K_a/K_s \) analysis demonstrates strong conservative selective pressure on both Zwilling proteins. However, the physiological role of Zwilling proteins for the function of myelinating glia of teleosts remains unclear. Fishes, unlike mammals and birds, are poikilothermic organisms that need to adapt their cellular machinery to varying temperatures. Membrane fluidity, the interactions of lipids and membrane-associated proteins, and therefore their impact on myelin membrane adhesion are temperature dependent and of crucial importance for proper myelin function (Hu et al. 2004). Although no direct sequence homology exists, Zwilling proteins resemble Myelin basic protein, one of the main myelin constituents, in terms of their physicochemical properties and their developmental expression profile. Zwilling proteins also possess a putative myristoyl moiety at their amino terminus, possibly targeting them to the myelin membrane. We hypothesize that Zwilling proteins may contribute to, or modulate the adhesion of the individual myelin lamellae in the teleost nervous system, perhaps by directly or indirectly modulating MBP membrane interactions. Specific protein kinases may thereby regulate through phosphorylation at phylogenetically conserved sites, the interaction of Zwilling proteins with the membrane, or other proteins such as MBP, in order to exert its physiological functions. Zwilling proteins are found in teleosts only. It is intriguing to speculate what molecular mechanisms other poikilothermic myelinating vertebrates such as cartilaginous and lobe-finned fishes, amphibia, and reptiles may have evolved to adjust to varying temperatures. With the sequencing of the genomes of more and more vertebrates of all classes, comparative evolutionary analyses may yield insights into these pathways.

With the identification of Zwilling proteins, as highly conserved teleost-specific components of myelinating glia, we have extended the list of myelin-specific proteins in teleosts. Future functional studies will likely yield interesting insights into the cellular role of Zwilling proteins, the myelin physiology of teleosts, and myelin evolution in general.

Methods

Zebrafish of the AB wildtype line were maintained and bred under standard conditions (Westerfield 1994). The full-length \( zwi \) cDNA was reconstructed, starting from EST AW281753, through alignments with EST sequences contained in Genbank, and using RT-PCR (Retroscript Kit, Ambion, Foster City, CA) and 5’-RACE-PCR (RLM-RACE Kit, Ambion, Foster City, CA) to bridge sequence gaps and determine the 5’ end of the transcript. MRNA in situ hybridization was carried out as described previously (Thisse et al. 1993) with EST clone EE210813 (a gift from Dr. H. Wang, University of Oklahoma) as template for cRNA probe synthesis. For biochemical analysis, zebrafish myelin membranes were prepared according to established protocols (Morris et al. 2004). Tryptic peptides with masses corresponding to ZwiA and ZwiB protein sequences were identified from SDS-PAGE gel cutsouts through matrix-assisted laser desorption ionization peptide mass fingerprint; for ZwiA these peptides were confirmed by peptide fragmentation fingerprint mass spectrometry (MS).

Protein Blast and PSI-Blast searches were conducted at the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify homologous sequences and the InterPro database (http://www.ebi.ac.uk/Tools/InterProScan/, European Bioinformatics Institute) was searched for conserved domains and patterns. Secondary structure prediction was carried out using the Chou−Fasman, GOR4, and PELE (which includes seven different algorithms) prediction methods implemented on the Biology WorkBench web server (http://workbench.sdsc.edu/). Myristoylation site prediction was performed on the MYR Prediction Server (mendel.imp.ac.at/myristate; Research Institute of Molecular Pathology, Vienna) and on the Prosite database web server (http://www.expasy.ch/prosite, Swiss Institute of Bioinformatics). Prosite was also used to detect phylogenetically conserved putative phosphorylation sites. Protein alignments were generated using ClustalW and phylogenetic trees constructed by DRAWTREE software, both of the PHYLIP phylogeny inference software package (Department of Genetics, University of Washington, Seattle). \( K_a/K_s \) values were computed with the \( K_a/K_s \) Calculation Tool of the Bergen Center for Computational Science (http://services.ebi.uib.no/tools/kaks). Genomic sequences around the \( zwi \) locus were obtained from the Sanger Centre’s Ensembl web site and used to assess conserved synteny and to generate percent–identity plots using MultiPiMaker software (Penn State University Center for Comparative Genomics and Bioinformatics, http://bio.cse.psu.edu/pipmaker).

Acknowledgments

We thank Drs. Lars Israel and Axel Imhof of the SFB596 Protein Analysis Unit for mass spectroscopy and Drs. Andrea Huber and Andrew Schaefer for critically reading the manuscript. This work was supported by a grant of the German Research Foundation (DFG SFB596/N01) to CB.

Literature Cited


Claudia Kappen, Associate Editor

Accepted December 22, 2008