Biol. Chem. Hoppe-Seyler
Vol. 372, pp. 865–874, September 1991

Evolution of the Myelin Integral Membrane Proteins of the Central Nervous System

Freimut Schliess and Wilhelm Stoffel

Institut für Biochemie, Medizinische Fakultät der Universität zu Köln

(Received 19 July 1991)

Summary: The predominant integral membrane protein of the CNS myelin of amphibia, reptiles, birds and mammals is proteolipid protein (PLP) and P_0, the main glycoprotein in PNS myelin. Alternative splicing of the transcripts of the single genes of PLP and myelin basic protein (MBP) is the underlying mechanism by which the isoforms of the two main proteins of the myelin membrane arise. DM20 is an isoform of PLP in mammalian, avian and reptilian myelin. It does not occur in the CNS myelin of amphibia. DM20 lacks an extended hydrophilic sequence exposed on the extracytoplasmic surface of the lipid bilayer as a result of the usage of a cryptic donor splice site within exon III.

We report about comparative studies on PLP and its DM20 isoform on the protein and DNA level of frog, chicken, rat CNS and the P_0-related IP proteins of the CNS of trout. Chemical cleavage at tryptophan residues with *N*-chlorosuccinimide yields identical patterns of PLP peptides which refers to a high conservation between amphibia, birds and mammals and is totally different from the cleavage pattern of hydrophobic myelin proteins IP-1 and IP-2 of trout CNS and that of P_0 of rat PNS.

The N-terminal 19 amino-acid residues of IP-1 of trout CNS- and P_0 of frog PNS myelin were sequenced and proved to be homologous on one hand with the P_0 analogue of CNS of the shark, a cartilage fish, and on the other hand with P_0 protein of PNS of birds and mammals.

The complete amino-acid sequence of chicken CNS PLP was derived from its cDNA. Coding and noncoding segments of the PLP gene of frog were sequenced: there is a high degree of conservation between amphibian and mammalian PLP within the hydrophobic domains. Numerous mutations were found within the part of exon III encoding the hydrophilic domain. Base exchanges within the putative splice site in exon III explain the absence of DM20 in the protein pattern of amphibia CNS myelin. This result is being discussed in view of the membrane organization and the function of PLP.

Zur Evolution der integralen Membranproteine des zentralen Nervensystems

Zusammenfassung: Das vorherrschende integrale Membranprotein des CNS-Myelins der Säugetiere, Vögel, Reptilien und Amphibien ist das Proteolipidprotein (PLP), im PNS-Myelin das P_0-Glycoprotein. Alternatives Spleißen der Transkripte des jeweils einmal vorliegenden Gens des PLP und des basischen

---

Enzyme:
Taq, Deoxynucleoside-triphosphate:DNA deoxynucleotidyltransferase, E.C. 2.7.7.7

Abbreviations:
CNS, central nervous system; DM20, PLP isoform with 241 amino-acid residues; MBP, myelin basic protein, a peripheral myelin protein with molecular mass 18.4 kDa; IP, integral membrane proteins of trout CNS myelin; PAGE, polyacrylamide gel electrophoresis; PLP, proteolipid protein of CNS myelin, a 276 amino-acid residues long integral membrane protein with molecular mass 31 kDa; PNS, peripheral nervous system; SDS, sodium dodecyl sulfate.

Myelinproteins (MBP) ist der Mechanismus, der zu den Isoformen der beiden wichtigsten Proteine der Myelinmembran führt. DM20 ist die Isoform des PLP im CNS-Myelin von Säugetieren, Vögeln und Reptilien, während es im CNS-Myelin der Amphibien nicht vorhanden ist. Im DM20 fehlt ein längerer hydrophiler Sequenzbereich, der zur extrazytosolischen Oberfläche der Lipiddoppelschicht gerichtet ist. Durch die Aktivierung einer kryptischen Spleiß-Donor-Stelle innerhalb von Exon III kommt es zu seiner Deletion.

Wir berichten über vergleichende Studien auf der Protein- und der DNA-Ebene zwischen dem PLP von Frosch-, Huhn- und Ratten-CNS sowie den P_0-verwandten IP-Proteinen des CNS der Forelle. Chemische Spaltung hinter Tryptophan mittels *N*-Chlorsuccinimid ergab identische Verteilungsmuster von PLP-Peptiden, die auf eine hohe Konservierung zwischen Amphibien, Vögeln und Säugetieren schließen lassen und völlig verschieden vom Peptidmuster der hydrophoben Myelinproteine IP-1 und IP-2 des Fisch-CNS und P_0 des Ratten-PNS sind.

Die *N*-terminalen 19 Aminosäuren von IP-1 aus Forellen-CNS-Myelin und von P_0 aus Frosch-PNS-Myelin wurden sequenziert und erwiesen sich als homolog zum P_0-Analogon des CNS vom Hai, eines Knorpelfisches, und zum P_0-Protein des PNS von Vögeln und Säugetieren.

Die vollständige Aminosäuresequenz des Hühner-PLPs wurde aus der cDNA abgeleitet. Kodierende und nichtkodierende Bereiche des PLP-Gens vom Frosch wurden sequenziert: eine hohe Konservierung innerhalb der hydrophoben Domänen wurden ermittelt. Zahlreiche Mutationen wurden innerhalb desjenigen Teils des Exons III gefunden, der die hydrophile Domäne kodiert. Basenaustausche an der hypothetischen Spleißstelle in Exon III erklären die Abwesenheit des DM20 im Proteinmuster des Amphibien CNS-Myelins. Diese Ergebnisse werden unter dem Gesichtspunkt der räumlichen Anordnung und der Funktion des PLP in der Myelinmembran diskutiert.

Key terms: PLP evolution, proteolipid protein isoforms, trout P_0 protein, frog PLP, chicken PLP.

Oligodendrocytes in CNS synthesize, assemble and redistribute the components of the myelin membrane into the processes of their plasma membrane which spirally wrap different axons with up to 100 multilayered segments separated by Ranvier's nodes. Myelination confers on these axons a rapid saltatory conductivity, allows the compact structures of CNS and reduces the energy needed for the repolarization of axonal membranes[1].

The advantageous structure of compact myelin is confined to chordata. Myelinated axons appear only in vertebrates from cartilage fish, bony fish, amphibia, reptiles, birds to mammals[2].

Two main protein components of CNS myelin are responsible for the compact myelin structure: myelin basic protein (MBP), a peripheral membrane protein, for the close apposition of the cytosolic surfaces of the plasma membrane processes of oligodendrocytes, and proteolipid protein (PLP), the integral membrane protein, for the apposition of the external surfaces of wrapping myelin processes due to its hydrophilic domains outside of the lipid bilayer. A group of MBP isoproteins is present in CNS and PNS of all species mentioned before. Proteolipid protein and P_0 protein show a different distribution over CNS and PNS in evolution[2]: CNS myelin of cartilage and bony fish and PNS myelin of all six vertebrate classes react with antibodies raised against P_0 glycoprotein of rat PNS. Antibodies against rat PLP recognize only antigens in CNS myelin of amphibia, reptiles, birds and mammals. PLP is therefore a recent protein in evolution with its first appearance in terrestrial tetrapods, different from P_0 and related proteins, which already occur in membrane structures of the nervous system of annelides[3] and represent the primitive progenitor molecule of immunoglobulins[4].

We here report studies on the integral myelin membrane proteins of CNS of members of four classes ascending in evolution: fish (trout), amphibia (frog), birds (chicken) and mammals (rat) a) on the protein level and b) on the DNA level.

Protein patterns of CNS myelin or its chloroform-methanol extract of trout, frog, chicken and rat show bands with similar apparent molecular masses. PLP and its isoprotein DM20 are present in rat and chicken. Frog CNS myelin contains only PLP, however, and is devoid of the DM20 isoprotein. The two bands in trout CNS myelin, IP-1 and IP-2, do not cross-react immunologically with anti rat PLP antibodies[5]. Comparison of the N-terminal amino-acid sequences

of IP-1 from the CNS of trout and P_0 from the PNS of frog, strongly supports the continuous development from fish P_0 analoga of central to P_0 of the peripheral nervous system of mammals. P_0 glycoprotein of CNS of fish is substituted by PLP in the evolutionary steps (amphibia), following bony fish.

Amphibia are unable to express DM20 protein[2]. Genomic DNA sequencing unveiled the molecular basis for the missing DM20: the coding sequence of the hydrophilic domain of exon III shows multiple mutations in the consensus sequence of the cryptic splice donor sequence in exon III, thereby prohibiting alternative splicing of the PLP transcript. Amino-acid exchanges of frog and chicken PLP accumulate in the extramembranal domains, not in the trans- and cis-membranal helices. This result and the possible impact for membrane stacking in myelin and the limits of evolution of PLP is discussed.

Materials and Methods

Animals

Frogs *(Xenopus laevis)* were obtained from Kähler (W-2000 Hamburg, Germany), rats from Lippische Versuchstierzucht Hagemann GmbH, W-4923 Extertal, Germany. Trouts *(Salmo gairdneri)* and chicken heads were purchased from local suppliers. The cDNA λgt10 library of 40-days-old chicken brain (male) was purchased from Clontech.

Myelin was prepared using a sucrose gradient according to Norton and Poduslo[6]. The hydrophobic membrane proteins were accumulated by chloroform/methanol (2:1) extraction. Electrophoretic separations were carried out after reductive carboxymethylation using the PAGE system of Laemmli[7] (15% acrylamide).

Tryptophan cleavage: the hydrophobic membrane proteins were cut out from the slab gel and incubated in a reaction system containing 15mM *N*-chlorosuccinimide in urea/acetic acid buffer[8]. After equilibration to appropriate conditions for SDS gel electrophoresis the peptides were separated on a second gel system (20% acrylamide) and made visible by silver staining[9].

Protein sequencing was carried out using automated Edman degradation (Applied Biosystems 477 A Protein Sequencer).

Polymerase chain reactions of DNA sequences of frog PLP were amplified using a DNA Thermal Cycler (Perkin Elmer, Norwalk, USA)[10].

Oligonucleotides used as sense and antisense primers were synthesized by phosphoramidite chemistry using the Applied Biosystems Model 380A automated synthesizer according to the Applied Biosystems directions. The oligonucleotides were desalted by chromatography on Sepak as recommended by the manufacturer. The resulting preparation was used without further purification. The oligonucleotides were derived from the rat PLP DNA sequence or, where indicated, derived from the results of the frog DNA sequence. Preparation of genomic frog DNA and blunt end cloning into the *Sma* I site of the polylinker sequence of the pUC13 vector and the screening of the chicken cDNA λgt10 library were carried out according to Maniatis *et al.*[11].

Nucleotide sequences were determined by double strand sequencing[12,13] with the chain termination method[14].

Results

1) Protein-chemical analysis of CNS myelin of hydrophobic proteins of fish, amphibia, birds and mammalia

Myelin was prepared[6] from trout, frog, chicken and rat brains. Aliquots were extracted with chloroform-methanol for PAGE[7] analysis.

Fig. 1 presents the SDS polyacrylamide gel electrophoretic separation of total CNS myelin proteins, lanes 1–4, and those present in the chloroform-methanol extract, lanes 5 and 6.

The strong band at 29 kDa in frog, chicken and rat myelin corresponds to PLP and the faster running band to the DM20 isoprotein in which amino-acid residues 116–150 are deleted. The bars indicate their relative position. In the chloroform-methanol extracts the hydrophobic proteins PLP and DM20 are accumulated. The extract of frog brain contains only PLP. The protein patterns of trout, chicken and rat are very similar, but the two bands with apparent molecular masses of PLP and DM20 in trout CNS myelin show no immunological relationship to PLP. They are referred to as IP-1 and IP-2[2] (bars in Fig. 1).

The purified 29 kDa proteins were chemically cleaved with *N*-chlorosuccinimide[8] at tryptophan residues Fig. 2.

The peptide patterns of the bovine, chicken and frog brain CNS PLP, although different in intensity, are very similar (lanes 1–3) and resemble that of earlier studies[15], but differed completely from that of the trout. IP-1 and IP-2 (lanes 5 and 6) of the trout brain yielded tryptophan fragment patterns which compared well with that of P_0 of PNS of rat (lane 4). The three principal bands of the cleavage pattern of IP-1 and IP-2 (lanes 5 and 6) are shifted in their electrophoretic mobility in parallel, which could be due to the different degree of glycosylation of the peptides.

We sequenced the purified IP-1 of trout CNS and P_0 of frog PNS and compared the sequence of the N-terminal 19 amino acids of both with known sequences of P_0 of bovine[16], rat[17] and chicken[18] peripheral and shark[19] central nervous system, Fig. 3.

Most of the amino-acid substitutions are conservative, some residues seem to be essential, e.g. Ile^1, Thr^5, Leu^{19} and the tripeptide Val^{13} to Ser^{15}. A continuous development of P_0 of the CNS of the shark to P_0 of the PNS of mammals and thereby a relationship between myelin of PNS of mammals and CNS of fish is apparent. The integral myelin proteins IP in CNS of the trout are replaced in frog, chicken and human by proteolipid protein with no homologies to P_0 glycoprotein.

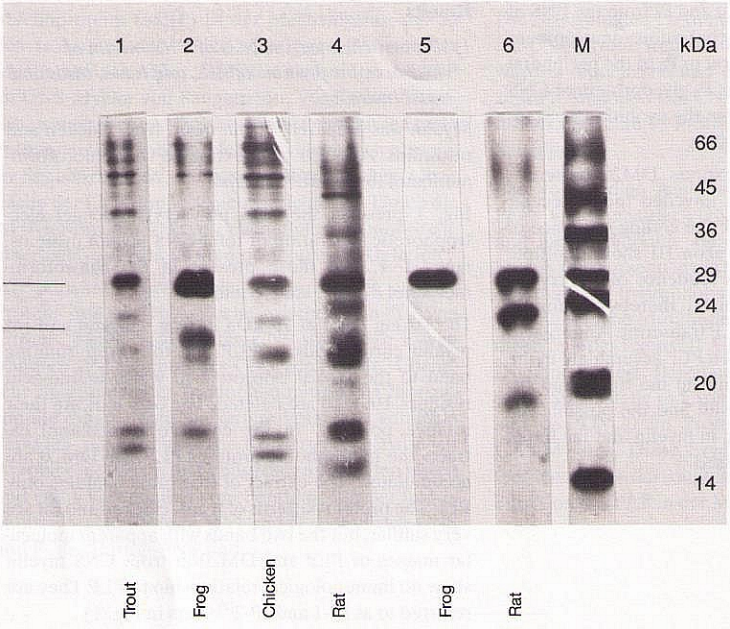

Fig. 1. Protein pattern of CNS myelin, lanes 1–4, of trout, frog, chicken and rat, and chloroform-methanol extracts of myelin of frog and rat brains, lanes 5 and 6.

M = Marker proteins, SDS PAGE 15%, silver staining. The bars indicate the position of the hydrophobic membrane proteins.

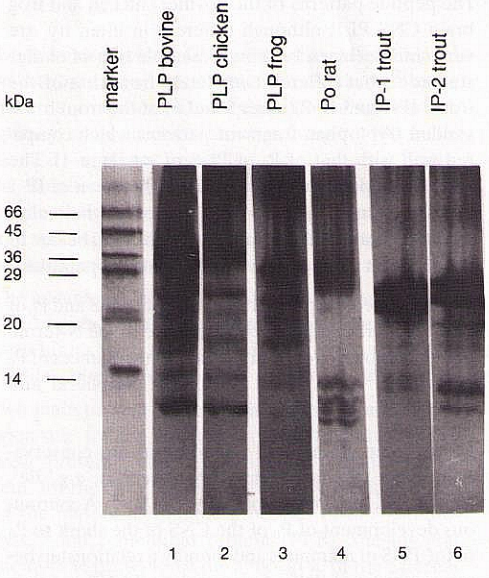

Fig. 2. Pattern of *N*-chlorosuccinimide cleavage of PLP from bovine, chicken and frog CNS, P_0 of rat PNS and IP-1 and IP-2 of trout CNS.

SDS PAGE, 20%, silver staining.

bovine	PNS	Ile	Val	Val	Tyr	Thr	Asp	Lys	Glu	Val	His	Gly	Ala	Val	Gly	Ser	Gln	Val	Thr	Leu
rat	PNS	Ile	Val	Val	Tyr	Thr	Asp	Arg	Glu	Val	Tyr	Gly	Ala	Val	Gly	Ser	Gln	Val	Thr	Leu
chicken	PNS	Ile	His	Val	Tyr	Thr	Pro	Arg	Glu	Val	Tyr	Gly	Thr	Val	Gly	Ser	His	Val	Thr	Leu
frog	PNS	Ile	Asp	Val	Tyr	Thr	Asp	Lys	Glu	Val	Tyr	Gly	Thr	Val	Gly	Ser	X	Val	Thr	Leu
trout	CNS	Ile	Val	Ile	Tyr	Thr	Gly	Trp	Glu	Leu	His	Ala	Leu	Val	Gly	Ser	Asp	Ile	Ile	Leu
shark	CNS	Ile	Ser	Val	Ser	Thr	His	His	Asn	Leu	His	Lys	Thr	Val	Gly	Ser	Asp	Val	Thr	Leu

Fig. 3. Comparison of amino-acid sequences of IP-1 of trout CNS, P_0 of frog PNS with P_0 of bovine, rat, chicken PNS and the shark CNS P_0 analogon.

The boxes indicate homologous residues.

2) Characterization of a complete chicken PLP cDNA clone

A λgt10 cDNA library of chicken brain was screened with a 700 bp ^{32}P-labelled *Pst* I fragment of the rat cDNA clone described before[20]. A 3-kb clone encoding the complete amino-acid sequence of PLP was isolated and sequenced with internal PLP primers. Fig. 4 documents the cDNA sequence of chicken PLP and the derived amino-acid sequence with an open reading frame of 891 bp encoding 277 amino-acid residues. These sequences are compared with the human PLP nucleotide and amino-acid sequence.

Chicken PLP is highly conserved on the amino acid level with few mostly conserved amino-acid exchanges (Val^{11}→Ile, Lys^{104}→Arg) as compared to the human sequence. A domain of frequent amino-acid exchanges is the hydrophilic domain encoded by exons III/IV which according to our proposed model of membrane integration forms an extracellular loop. This is in accordance with the data described for this domain of the amphibian PLP.

Contrary to exon III of frog PLP the chicken sequence has developed the cryptic splicing site (nucleotides 354–397 downstream initiation codon ATG) leading to the DM20 PLP isoform during expression.

3) Mutations in coding sequences of exon III of frog lead to the DM20 isoform of proteolipid protein

The direct access to the molecular basis of the appearance of the DM20 isoform of PLP during evolution came from the comparative sequence analysis within exons III, intron III and exon IV of frog and human[21]. The protein pattern of frog CNS myelin is devoid of the DM20 protein band in PAGE (Fig. 1, lane 5). Parts of the PLP gene of frog were analysed by PCR[10] using synthetic oligonucleotide primers homologous to strongly conserved domains of exons III and IV. Fig. 5 lists the nucleotide sequence of exon III and large parts of intron II and exon IV of the human and frog PLP gene. The considerable deviation of the frog sequence is striking. The potential alternate 5′ donor splice site within exon III of mammals, responsible for the deletion of the 115 bp at the 3′ end of exon III leading to DM20 is encased.

The 5′ donor splice site of exon III–intron III and that of the intra exon III site are compared with the consensus sequence AG/GT $\frac{A}{G}$ AGT[22] and with the corresponding exon III sequence of the frog and chicken PLP gene (see table).

No consensus can be seen 5′ of the GT sequence within exon III of frog. By comparing the intron sequences of human and frog it is found that the homology is strongly reduced. The bases of intron/exon transitions (AG and GT, respectively) are identical between frog and mammals. The amino-acid sequence derived from the coding nucleotide sequences within exons III and IV is shown in Fig. 6.

There are two residues exchanged in the hydrophobic transmembrane domain, one of which is conservative (Ala^{83}→Ile). Amino acid substitutions, which are framed in Fig. 6, accumulate in the extramembranal hydrophilic loop of the model for the PLP integration into the lipid bilayer[23]. Two of the three cysteine residues (Cys^{138} and Cys^{140}) are conserved in frog PLP, which strongly indicates their contribution to essential functions of the hydrophilic loops on the extracytosolic surface. Cys^{118} is exchanged against Met. A proline residue is inserted between Gln^{134} and Lys^{135} in frog PLP. An inversion of Lys^{122}-Gly^{123} in mammalian PLP has occurred in the frog PLP. Otherwise it is striking to notice the conserved number of charged side chains, two negative and eight positive in mammals versus two negative and nine positive in amphibia, and in addition their conserved distribution along the polypeptide chain.

```
     g   t   aat tg  agt cag agt c c aaa  a          T    A              A A A
-30  cgc aac ggg gag ctg agc gag tcg ggt gcc ATG GGT CTG TTG GAG TGC TGT GCC CGC TGT   30
  0                                        Met Gly Leu Leu Glu Cys Cys Ala Arg Cys

       G G         C       T  T  C      G          A T G  T                G  A
 31  CTC ATA GGG GCA CCC TTC GCC TCT CTG GTC GCC ACT GGG CTC TGC TTC TTT GGG GTC GCG   90
 10  Leu Ile Gly Ala Pro Phe Ala Ser Leu Val Ala Thr Gly Leu Cys Phe Phe Gly Val Ala
        Val

               T       T  A  T              T       A  A A    A              T
 91  CTG TTC TGC GGC TGC GGG CAC GAA GCC CTC ACC GGC ACC GAG CAG CTC ATT GAG ACC TAC   150
 30  Leu Phe Cys Gly Cys Gly His Glu Ala Leu Thr Gly Thr Glu Gln Leu Ile Glu Thr Tyr
                                                          Lys

             A          A       T              C A    G       T  C  C       T
151  TTC TCC AAG AAC TAC CAG GAC TAC GAG TAT CTC ATT GAT GTC ATC CAC GCT TTT CAG TAC   210
 50  Phe Ser Lys Asn Tyr Gln Asp Tyr Glu Tyr Leu Ile Asp Val Ile His Ala Phe Gln Tyr
                                                    Asn

                   T       T              T  T  G       C              T  G
211  GTC ATC TAT GGA ACA GCC TCC TTC TTC TTC CTC TAC GGA GCC CTG CTG CTG GCC GAA GGC   270
 70  Val Ile Tyr Gly Thr Ala Ser Phe Phe Phe Leu Tyr Gly Ala Leu Leu Leu Ala Glu Gly

                              A    G       T  C          AA
271  TTC TAC ACC ACC GGC GCA GTC CGG CAA ATC TTC GGG GAC TAC CGG ACC ACC ATC TGC GGC   330
 90  Phe Tyr Thr Thr Gly Ala Val Arg Gln Ile Phe Gly Asp Tyr Arg Thr Thr Ile Cys Gly
                                                       Lys

             G                  A          A   G          T T C A       AA  T
331  AAG GGC CTC AGC GCA ACG GTA ACT GGG GGC CCG AAA GGG AGG GGA GCG CGA GGC CCC CAG   390
110  Lys Gly Leu Ser Ala Thr Val Thr Gly Gly Pro Lys Gly Arg Gly Ala Arg Gly Pro Gln
                                        Gln                Ser         Gln His

     A       T          G                    T                  A              C
391  CGA GCT CAC TCT TTG CAG CGG GTG TGT CAG TGT TTG GGA AAG TGG CTA GGA CAT CCT GAC   450
130  Arg Ala His Ser Leu Gln Arg Val Cys Gln Cys Leu Gly Lys Trp Leu Gly His Pro Asp
     Gln                Glu              His

                    C          C       T G T  A                TG  T          T
451  AAG TTT GTG GGC ATT ACC TAT GTC CTG ACC ATC GTC TGG CTC CTG GCC TTC GCC TGC TCC   510
150  Lys Phe Val Gly Ile Thr Tyr Val Leu Thr Ile Val Trp Leu Leu Ala Phe Ala Cys Ser
                           Ala          Val              Val

       T       T  G       T       C  T                          T  T          T
511  GCC GTG CCC GTC TAC ATC TAC TTT AAC ACC TGG ACC ACC TGC CAG TCC ATC GCC TTC CCA   570
170  Ala Val Pro Val Tyr Ile Tyr Phe Asn Thr Trp Thr Thr Cys Gln Ser Ile Ala Phe Pro

     G         T         T  A     GT  C       T  T       A       T       T  C
571  ACC AAG ACC ACT GCC AGC ATC GGC ACG CTG TGC GCG GAC GCC AGG ATG TAC GGT GTC CTG   630
190  Thr Lys Thr Thr Ala Ser Ile Gly Thr Leu Cys Ala Asp Ala Arg Met Tyr Gly Val Leu
     Ser         Ser                Ser

       A       T  T       T  C       T                      T  G                A  A
631  CCC TGG AAC GCG TTT CCC GGG AAG GTG TGT GGC TCC AAC CTG CTC TCC ATC TGC AAG ACC   690
210  Pro Trp Asn Ala Phe Pro Gly Lys Val Cys Gly Ser Asn Leu Leu Ser Ile Cys Lys Thr

     GC                    C          G  T  T  T  A          T       A       A
691  AGC GAG TTC CAA ATG ACT TTC CAC CTC TTC ATC GCG GCC TTT GTG GGG GCT GCC GCC ACT   750
230  Ser Glu Phe Gln Met Thr Phe His Leu Phe Ile Ala Ala Phe Val Gly Ala Ala Ala Thr
     Ala

       A  T  C                       T  T                   T           T  A  C
751  CTG GTC TCA CTG CTC ACC TTC ATG ATC GCC GCC ACT TAC AAC TTC GCC GTC CTC AAG CTG   810
250  Leu Val Ser Leu Leu Thr Phe Met Ile Ala Ala Thr Tyr Asn Phe Ala Val Leu Lys Leu

             A                  ga t c c c gta gaa t    c ctt tct  ta at  g g agg
811  ATG GGC CGG GGC ACC AAG TTC tag ccg gcg agg tgg acc cca gcc aga cac gca ccc tcc   870
270  Met Gly Arg Gly Thr Lys Phe

     c c taa  c  ca  agc cta  aa  gc tg  gt  tc    t   t  ac t t t
871  ttt cct cta acc ccg agg ctt taa cac tgc agc cac ctg aca cca g                    916
```

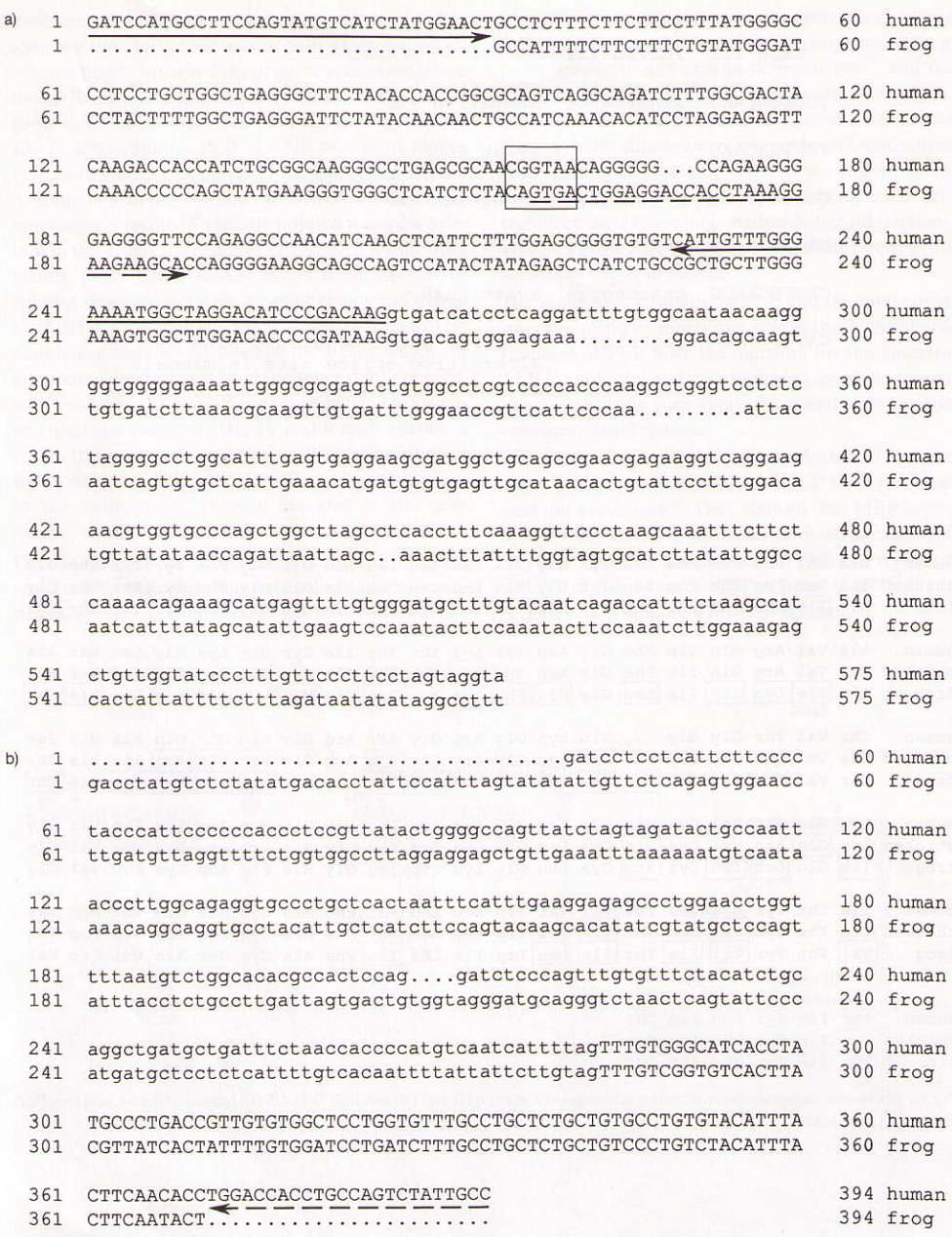

Fig. 5. Comparison of human and frog PLP nucleotide sequences of a) exon III, b) exon IV and the intervening intron. The arrows indicate sense and antisense primers of the two PCR reactions (first PCR: —, second PCR: – – –). The box shows the human alternative splice site leading to the DM20 isoform and the corresponding frog sequence.

◁ Fig. 4. Complete nucleotide sequence of chicken PLP cDNA and derived amino-acid sequence. Human nucleotide and amino-acid deviations are noted above and below the chicken sequence. The DM20 cryptic splice site is encased.

Table. Comparison of splice junction sites.

Exon III	Intron III	
TCCCGACAAG	gtgatcatcc	mammalian PLP
TCCTGACAAG	-	avian PLP
CCCTGATAAG	gtgacagtgg	amphibian PLP
GAGCGCAACG	gtaacagggg	mammalian DM20
CAGCGCAACG	gtaactgggg	avian DM20
CATCTCTACA	gtgactggag	frog PLP sequence around site of alternative splice site in mammals
AG	gtaagt	consensus 5' splice site

```
          75
human     Ala Ser Phe Phe Phe Leu Tyr Gly Ala Leu Leu Leu Ala Glu Gly Phe Tyr Thr Thr Gly
chicken   Ala Ser Phe Phe Phe Leu Tyr Gly Ala Leu Leu Leu Ala Glu Gly Phe Tyr Thr Thr Gly
frog      Ala [Ile] Phe Phe Phe Leu Tyr Gly [Ile] Leu Leu Leu Ala Glu Gly Phe Tyr Thr Thr [Thr]

human     Ala Val Arg Gln Ile Phe Gly Asp Tyr Lys Thr Thr Ile Cys Gly Lys Gly Leu Ser Ala
chicken   Ala Val Arg Gln Ile Phe Gly Asp Tyr [Arg] Thr Thr Ile Cys Gly Lys Gly Leu Ser Ala
frog      Ala [Ile] [Lys] [His] Ile [Leu] Gly [Glu] [Phe] Lys [Pro] [Pro] [Ala] [Met] [Lys] [Gly] Gly Leu [Ile] [Ser]

human     Thr Val Thr Gly Gly ... Gln Lys Gly Arg Gly Ser Arg Gly Gln His Gln Ala His Ser
chicken   Thr Val Thr Gly Gly ... [Pro] Lys Gly Arg Gly [Ala] Arg Gly [Pro] [Gln] [Arg] Ala His Ser
frog      Thr Val Thr Gly Gly [Pro] [Pro] Lys Gly Arg [Ser] [Thr] Arg Gly [Arg] [Gln] [Pro] [Val] His [Thr]

human     Leu Glu Arg Val Cys His Cys Leu Gly Lys Trp Leu Gly His Pro Asp Lys Phe Val Gly
chicken   Leu [Gln] Arg Val Cys [Gln] Cys Leu Gly Lys Trp Leu Gly His Pro Asp Lys Phe Val Gly
frog      [Ile] Glu [Leu] [Ile] Cys [Arg] Cys Leu Gly Lys Trp Leu Gly His Pro Asp Lys Phe Val Gly

human     Ile Thr Tyr Ala Leu Thr Val Val Trp Leu Leu Val Phe Ala Cys Ser Ala Val Pro Val
chicken   Ile Thr Tyr [Val] Leu Thr [Ile] Val Trp Leu Leu Val Phe Ala Cys Ser Ala Val Pro Val
frog      [Val] Thr Tyr [Val] [Ile] Thr [Ile] [Leu] Trp Ile Leu [Ile] Phe Ala Cys Ser Ala Val Pro Val

                        179
human     Tyr Ile Tyr Phe Asn Thr
chicken   Tyr Ile Tyr Phe Asn Thr
frog      Tyr Ile Tyr Phe Asn Thr
```

Fig. 6. Nucleotide sequence-derived amino-acid sequence of exon III and IV (residues 75 to 179) of human, chicken, and frog PLP. Exchanges are framed.

Within the amino acids of exon IV only conservative exchanges are being observed. Thr^{155} – its exchange to Ile causes an exon IV form of Pelizaeus Merzbacher disease[24] – is not changed within the frog and chicken sequence. The part of the sequence being responsible for the return of the α-helix and the formation of a cis-membranal loop[23] is completely conserved.

Discussion

The structural analysis of the main integral membrane protein PLP of human[25], bovine[26], rat[20], mouse[27] and dog[28] CNS myelin pointed out the extremely high homology of the 276 amino-acid residues of PLP of these species. The polypeptide is strictly structured into hydrophobic and hydrophilic domains, the di-

mensions of the hydrophobic sequences are suited to span the lipid bilayer of myelin with its special width as transmembrane α-helices or enter as cis-membrane loops[23]. PLP is encoded in 7 exons of the 17 kb gene[21]. The PLP locus has been assigned to the Xq 13–22 chromosome. PLP of CNS myelin of higher vertebrates is accompanied by the isoprotein DM20. Amphibians do not contain the DM20 protein[2]. Immunological results[29] initially pointed toward a deletion in the 110 to 160 residues domain. DNA analyses clearly proved that DM20 arises from alternative splicing in which a cryptic 5′ splice donor site within exon III is activated and leads to the deletion of 35 amino acid residues encoded by the 105 bp which are eliminated downstream from the splice donor consensus[30] sequence. The PLP gene is organized in such a way that the exons II, III, IV and V each encode a single trans- or cis-membrane hydrophobic domain followed by a hydrophilic domain as connecting link to the hydrophobic domain encoded in the next exon[21].

The meaning of the strong conservation of the PLP structure during evolution particularly of the mammalian species is underlined by the fatal consequence of point mutations within exon sequences encoding hydrophobic domains, leading to dysmyelinoses, e.g. jimpy mouse[27] and myelin-deficient rat[31] and the different cases of Pelizaeus-Merzbacher disease reported so far[24]. Mutations in animal models and human dysmyelinoses have not yet been found within the hydrophilic domains. The trans- and cis-membrane helices are apparently highly ordered and stabilized by H-bonding, hydrophobic interactions, aromatic stacking within and by disulfide bonds at the surface of the lipid bilayer.

The conserved structure of PLP and the fatal consequences of point mutations within the hydrophobic domains of PLP raise the question for the ancestral PLP in evolution and the modifications of the protein during evolution very likely associated with putative conformational changes.

Immunological studies on the evolution of the two main myelin proteins MBP and PLP have been carried out previously[2]. They showed that MBP is present in CNS and PNS of all classes down to cartilaginous fish. P_0 occurs ubiquitously in PNS and is present in CNS of fish, whereas PLP appears in CNS of vertebrates from amphibia through mammals.

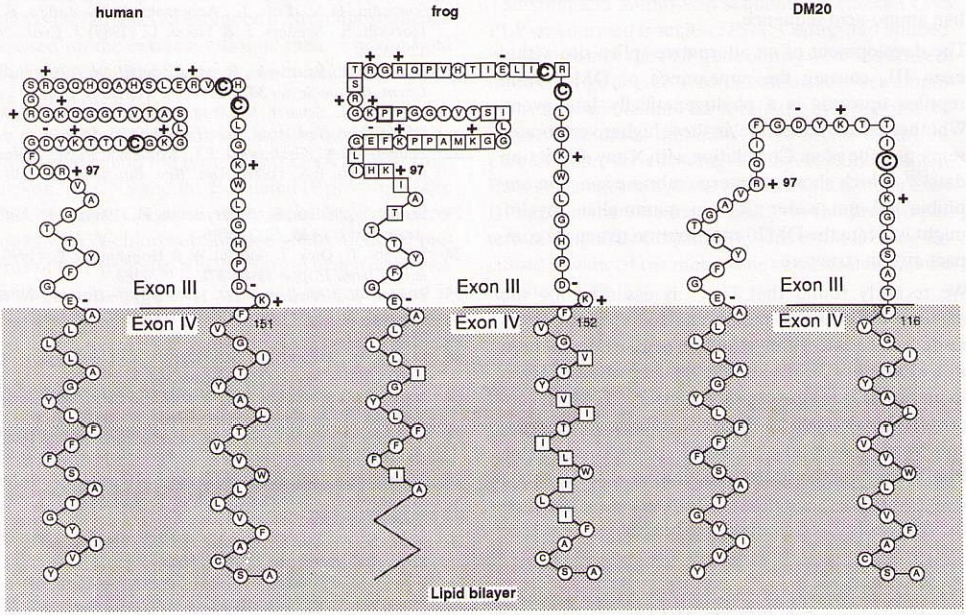

Fig. 7. Amino-acid sequence of the hydrophilic loop connecting trans- and cis-membranal α-helices encoded by exon III and IV, respectively with the distribution of charged residues.

Amino-acid exchanges: squares, cysteines: bold circles, proline insertion: bold square. a) Human PLP; b) frog PLP; c) DM20 isoform.

The protein sequence studies and recombinant DNA techniques reported here provide data on the molecular and genomic level and thereby allow a detailed analysis of the evolution of myelin proteins and give insight into their function during myelinogenesis of the central nervous system.

We could show, that most of the amino-acid exchanges presented here are conservative. The frequency of exchanges is higher in the extracellular loop (exon III), than in the trans-(exon III) or cis-(exon IV) membranal α-helical domains. The conserved positions of cysteines (e.g. Cys^{138}, Cys^{140}, Cys^{167}) indicate their involvement in structural important disulfide bridges.

The almost identical charge pattern in the hydrophilic loop, coded by exon III and orientated towards the extracellular space refers to a specific surface pattern, which might interact with complementary structures in the wrapping process and stacking of myelin layers, Fig. 7.

Previous immunological data[32] showed, that antibodies against a synthetic peptide (residues 116–128) located within the hydrophilic loop discussed here strongly react with the PLP of higher vertebrates whereas the reaction with PLP of amphibia is weak. Here we present the molecular basis by showing that a high number of amino-acid exchanges is present in the amphibia PLP gene as compared to the mammalian amino-acid sequence.

The development of an alternative splice site within exon III, causing the appearance of DM20 from reptiles upwards is a phylogenetically late event. Whether DM20 is essential in these higher vertebrates remains to be seen. Correlation with X-ray diffraction data[33], which show an intermembrane space in amphibia 0.5 nm wider than in mammalian myelin, might indicate the DM20 contribution to a more compact myelin structure.

We recently found that Thr^{155} is essential for the stabilization of the cis-membranal loop (exon IV). Its mutation leads to dysmyelination in man documented in a kindred of Pelizaeus-Merzbacher disease. Even in amphibia this Thr residue is conserved, which supports its importance for the protein function. Also the complete conservation of the peptide sequence Phe^{165}-Thr^{180} within exon IV is remarkable. We have assigned to this sequence a β-turn structure between the descending and ascending α-helices.

The generous support of these studies by the *Deutsche Forschungsgemeinschaft*, SFB 243, Molekulare Analyse und Entwicklung zellulärer Systeme, and the *Fritz-Thyssen-Stiftung* is gratefully acknowledged.

References

1 Raine, C.S. (1984) in *Myelin*, (Morell, P., eds.) pp. 1–50, Plenum Press, New York.
2 Waehneldt, T., Matthieu, J. & Jeserich, G. (1986) *Neurochem. Int.* **9**, 463–474.
3 Pereyra, P.M. & Roots, B.I. (1988) *Neurochem. Res.* **13**, 893–901.
4 Lemke, G., Lamar, E. & Patterson, J. (1988) *Neuron* **1**, 73–83.
5 Jeserich, G. & Waehneldt, T.V. (1987) *Neurochem. Res.* **9**, 825–829.
6 Norton, W. & Poduslo, S. (1973) *J. Neurochem.* **21**, 749–757.
7 Laemmli, U.K. (1970) *Nature (London)* **227**, 680–685.
8 Lischwe, M.A. & Ochs, D. (1982) *Anal. Biochem.* **127**, 453–457.
9 Blum, H., Beier, H. & Gross, H.J. (1987) *Electrophoresis* **8**, 93–99.
10 Saiki, R., Gelfand, D., Stoffel, S., Scharf, S., Higuchi, R., Horn, G., Mullis, K. & Erlich, H. (1988) *Science* **239**, 487–491.
11 Maniatis, T., Fritsch, E.F. & Sambrock, J. (1982) *Molecular Cloning*, Cold Spring Harbor, New York.
12 Chen, E.J. & Seeburg, P.H. (1985) *DNA* **4**, 165–170.
13 Murphy, G. & Kavanagh, T. (1988) *Nucleic Acids Research* **16**, 5198.
14 Sanger, F., Nicklen, S. & Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
15 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1982) *Hoppe-Seyler's Zeitschr. Physiol. Chemie* **363**, 1397–1407.
16 Sakamoto, Y., Kitamura, K., Yoshimura, K., Nishijama, T. & Uyemura, K. (1987) *J. Biol. Chem.* **25**, 4208–4214.
17 Lemke, G. & Axel, R. (1985) *Cell* **40**, 501–508.
18 Barbu, M. (1990) *J. Neurosc. Res.* **25**, 143–151.
19 Saavedra, R.A., Fors, L., Aebersold, R.H., Arden, B., Horvath, S., Sanders, J. & Hood, L. (1989) *J. Evol.* **29**, 129–156.
20 Schaich, M., Budzinski, R.-M. & Stoffel, W. (1986) *Biol. Chem. Hoppe-Seyler* **367**, 825–834.
21 Diehl, H.-J., Schaich, M., Budzinski, R.-M. & Stoffel, W. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9807–9811.
22 Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. & Sharp, P.A. (1986) *Ann. Rev. Biochem.* **55**, 1119–1150.
23 Stoffel, W., Hillen, H. & Giersiefen, H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5012–5016.
24 Weimbs, T., Dick, T., Stoffel, W. & Boltshauser, E. (1990) *Biol. Chem. Hoppe-Seyler* **371**, 1175–1183.
25 Stoffel, W., Giersiefen, H., Hillen, H., Schröder, W. & Tunggal, B. (1985) *Biol. Chem. Hoppe-Seyler* **366**, 627–635.
26 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1983) *Biol. Chem. Hoppe-Seyler* **364**, 1455–1466.
27 Nave, K., Lai, C., Bloom, F. & Milner, R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9264–9268.
28 Nadon, N.L., Duncan, I.D. & Hudson, L.D. (1990) *Development* **110**, 529–537.
29 Trifilieff, E., Luu, B., Nussbaum, J., Roussel, G., de los Monteros, E., Sabatier, J. & Van Rietschoten, J. (1986) *FEBS Lett.* **198**, 235–239.
30 Nave, K.-A., Lai, C., Bloom, F.E. & Milner, R.J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5665–5669.
31 Boison, D. & Stoffel, W. (1989) *EMBO-J.* **8**, 3295–3302.
32 Linington, C. & Waehneldt, T.V. (1990) *J. Neurochem.* **54**, 1354–1359.
33 Kirschner, D., Inouye, H., Ganser, A. & Mann, V. (1989) *J. Neurochem.* **53**, 1599–1609.

Freimut Schließ and Wilhelm Stoffel, Institut für Biochemie, Medizinische Fakultät der Universität zu Köln, Joseph-Stelzmann-Straße 52, D-5000 Köln 41.