Biol. Chem. Hoppe-Seyler
Vol. 366, pp. 627–635, July 1985

Amino-Acid Sequence of Human and Bovine Brain Myelin Proteolipid Protein (Lipophilin) is Completely Conserved

Wilhelm STOFFEL, Helmut GIERSIEFEN, Heinz HILLEN, Werner SCHROEDER and Budi TUNGGAL

Institut für Physiologische Chemie der Universität zu Köln

(Received 1 April/28 May 1985)

Summary: Proteolipid protein (PLP) was isolated from white matter of human brain by chloroform/methanol extraction and further purified by chromatography. Performic acid oxidation yielded a product homogenous in $NaDodSO_4$-polyacrylamide electrophoresis with a molecular mass of 30 kDa. The carboxymethylated PLP was chemically cleaved with cyanogen bromide into four fragments: CNBr I 22–24 kDa, CNBr II 5 kDa, CNBr III 1.4 kDa and CNBr IV 0.7 kDa. HBr/dimethylsulfoxide cleavage at tryptophan residues released four fragments: Trp I 14–16 kDa, Trp II 2.0 kDa, Trp III 5 kDa and Trp IV 7 kDa. Hydrophilic fragments were enriched in 50% formic acid (CNBr II, III, IV and Trp II and III), whereas hydrophobic peptides precipitated from this solvent were CNBr I, Trp I and IV. The fragments were separated by gel filtration with 90% formic acid as solvent and finally purified by gel permeation HPLC (Si 60 and Si 100) for automated liquid and solid-phase Edman degradation. Large fragments were further cleaved with different proteinases (trypsin, V8-proteinase, endoproteinase Lys-C and thermolysin). We used an improved strategy in the sequencing of the human proteolipid protein compared with our approach to the structural elucidation of bovine brain PLP.
The amino-acid sequence of human PLP contains 276 residues, the same as found in bovine proteolipid protein. The two sequences proved to be identical. The possible importance of the conservative structure of this integral membrane protein is discussed.

Die Aminosäure-Sequenz des Proteolipidproteins (Lipophilin) des Menschen- und Rindergehirn-Myelins ist vollständig konserviert

Zusammenfassung: Proteolipid-Protein (PLP) wurde aus weißer Substanz von menschlichem Gehirn mit Chloroform/Methanol extrahiert und durch Chromatographie gereinigt. Das Protein erwies sich nach Perameisensäure-Oxidation in der $NaDodSO_4$-Polyacrylamidgel-Elektrophorese einheitlich und hatte eine molekulare Masse von 30 kDa. Das carboxymethylierte PLP wurde sowohl der Bromcyanspaltung an Methioninresten als auch der HBr/Dimethylsulfoxid-Spaltung an Tryptophanresten der Polypeptidkette unterworfen. Hydrophobe Fragmente (CNBr I und Trp I and IV) wurden in 50proz. Ameisensäure ausgefällt, über Gelfiltration in 90proz. Ameisensäure getrennt und durch wiederholte Gelpermeation-HPLC (Si 60 und Si 100) für die Sequenzierung durch automatischen Flüssig- und Festphasen-Edman-Abbau gereinigt. Vier Bromcyanfragmente CNBr I 22–24 kDa, II 5 kDa, III 1.4 kDa und IV 0.7 kDa sowie vier Tryptophanbruchstücke Trp I 14–16 kDa, II 2.0 kDa, III 5 kDa und IV 7 kDa wurden isoliert. Größere Fragmente wurden weiter proteolytisch durch Trypsin, V8-Proteinase, Endoproteinase Lys-C und Thermolysin abgebaut.
Für die Sequenzierung des menschlichen PLP wurde eine im Vergleich zu der von uns bei der Strukturaufklärung des Rinder PLP angewandten Methodik verbesserte Sequenzierungs-Strategie verfolgt. Die Aminosäuresequenz des mensch-

---

Abbreviations: PLP = proteolipid protein, lipophilin; PTH = phenylthiohydantoin; MBP = myelin basic protein; HPLC = high performance liquid chromatography; $NaDodSO_4$ = sodium dodecyl sulfate.

lichen PLP enthielt 276 Reste gleich der des Rindergehirn PLP. Die beiden Sequenzen waren identisch. Die mögliche Bedeutung der konservativen Struktur des 276 Aminosäurereste enthaltenden integralen Membranproteins wird diskutiert.

Key words: Lipophilin, human and bovine myelin, conserved amino-acid sequence, trans-membranous protein, phylogeny.

The multilamellar membrane system of myelin in the central nervous system is formed by oligodendrocytes as extension of their plasma membranes with a highly specialized lipid and protein composition. The plasma membrane protrusions wrap around the axon spirally forming morphologically compact, close appositions of the cytoplasmic faces of the membrane, visible as main dense line of the trilaminar periodicity of myelin in electron microscopy. The intermediate dense line corresponds to the tight apposition of the external plasma membrane faces[1–3].

Functionally this membrane system guarantees the insulator properties required for the unperturbed saltatory transmission of the nerve impulse along the nerve cell axons. One extrinsic, water-soluble protein, the myelin basic protein (MBP) and one integral, extremely hydrophobic protein, proteolipid protein (PLP) or lipophilin, are the main protein components of myelin, amounting to 90%. MBP, rich in basic and polar amino-acid residues has a molecular mass of 18 kDa. Its structure has been elucidated by the regular analytical techniques for the determination of water-soluble protein structures[4,5].

Recently the amino-acid sequence of bovine myelin proteolipid protein has been elucidated in this laboratory[6–8]. The protein containing 276 amino-acid residues is characterized by a hydrophilic N-terminus and C-terminus and four additional hydrophilic sequences which separate five hydrophobic domains containing long stretches of uncharged residues. They are bordered by single positively or negatively charged or by pairs of oppositely charged residues. At least 10 of the 16 cysteine residues form disulfide bridges.

On the basis of our sequence data we have constructed a model for the integration of proteolipid protein in the membrane[9]. We proposed that three of the five hydrophobic sequences are transmembrane segments spanning the lipid bilayer. We aligned two hydrophobic sequences as cis-domains, which enter and emerge on the same side of the membrane. They each contain a proline residue around the center of the sequence which induces a β-turn.

This hypothetical model has been experimentally supported by trypsin attack on osmotically dissociated rat myelin multilayer membranes[9]. The sparse comparative sequence data on rat, bovine and human myelin PLP indicated homology in the amino-acid sequence[11].

In this communication we describe the complete amino-acid sequence of proteolipid protein isolated and purified from white matter of human brain. PLP of human origin contains 276 amino-acid residues like that of bovine origin. The two protein sequences are completely conserved; no amino-acid residue is altered.

The structure of human lipophilin was determined with established and improved separation and sequencing techniques for hydrophobic polypeptides, used successfully in our structural studies on bovine myelin PLP[6–8].

Materials and Methods

Proteolipid protein was extracted from the white matter of human brain with chloroform/methanol by established procedures[12]. It was further purified by Sephadex LH-60 chromatography[6]. 0.8 to 1 g of PLP could be obtained from one brain.

Trypsin, thermolysin, V8-proteinase and endoproteinase Lys-C (from *Lysobacter enzymogenes*) were purchased from Boehringer, Mannheim.

Performic acid oxidation. 5 ml of 5% H_2O_2 solution in 98% formic acid was added dropwise to 100 mg PLP dissolved in 10 ml formic acid/methanol (v/v 9 : 1) at − 15 °C. The reaction was completed within 4 h. Solvents were removed by rotary distillation at room temperature.

Maleoylation of PLP. 200 mg reduced and carboxymethylated PLP[13] was dissolved in 1% $NaDodSO_4$ and 2 g of maleic anhydride added in small aliquot portions. The pH was kept constant between pH 8.5 to 9.5 by the addition of 2M NaOH, the reaction continued for 1 h at room temperature after the addition of maleic anhydride. The solution was dialysed against 10mM NH_4HCO_3.

Cyanogenbromide cleavage was performed as described for bovine PLP[14].

HBr-dimethyl sulfoxide cleavage[15] was adapted to the cleavage of the hydrophobic PLP: 150 mg carboxymethylated PLP was dissolved in 3 ml trifluoroacetic acid and 3 ml 12M HCl.

400 µl dimethyl sulfoxide was added and the solution stirred at room temperature for 30 min. After the addition of 400 µl dimethyl sulfoxide and 3 ml 48% HBr the solution was stirred for additional 30 min. The protein was precipitated by the addition of 10 ml H_2O and isolated by centrifugation at 200 × *g* for 15 min. The sediment was washed with acetone, air-dried and then dissolved in 90% formic acid for complete cleavage at 40–50 °C for 24 h. The Trp^{180}-Thr^{181} linkage in Trp-fragment III was split only after prolonged formic acid treatment at this elevated temperature.

Proteolytic cleavages

Tryptic digestion. 200 mg reduced and carboxymethylated PLP, dissolved in 50 ml 0.1M NH_4HCO_3 buffer, pH 8.6 was incubated with 4 mg trypsin at 37 °C for 18 h. The reaction was stopped by acidification with 6M HCl. The precipitate was dissolved in formic acid for chromatography.

Thermosinolytic cleavage was carried out in 50 ml 0.1M ammonium acetate buffer pH 7.7 for 3 h at 37 °C, protein/thermolysin 100 : 1 (w/w).

V8-proteinase treatment of Trp I. 10 mg Trp I in 500 µl trifluoroacetic acid was dried in a nitrogen stream, the peptide dissolved in 1 ml 0.5M NaOH and after addition of 3 ml water, 2 ml 0.1% $NaDodSO_4$, 200 µl 0.2M EDTA, the pH was adjusted to 7.9 with 2 ml 1M $NaHCO_3$. 1 mg V8-proteinase was added and the solution stirred at 37 °C for 24 h.

Endoproteinase Lys-C digestion. 100 mg reduced and carboxymethylated PLP was digested in 50 ml 0.1M NH_4HCO_3 buffer pH 8.6 with 1 mg endoproteinase Lys-C for 12 h at 37 °C with constant stirring.

The endoproteinase Lys-C fragment was isolated from the peptide mixture by gel permeation HPLC. A peptide with a molecular mass of 6 kDa was purified for sequencing. The conditions for solvent fractionation with formic acid/water mixtures, gel filtration (Biogel P10, P60, P150, gel permeation HPLC (Si 60, Si 100 columns) and liquid phase automated Edman degradation have been described previously[6–8].

For solid-phase automated Edman degradation polypeptides were coupled to diisothiocyanate glass in non-aqueous medium[16]. A home-made automatic solid phase sequenator was used[17].

Results

Sequence data from fragments of the cyanogenbromide cleavage in 90% formic acid of purified PLP yielded four peptides: CNBr I 22 kDa, CNBr II 5 kDa, CNBr III 1.4 kDa and CNBr IV 0.7 kDa. Molecular masses were derived from standardized gel permeation HPLC[6]. CNBr I was precipitated from 50% formic acid, whereas CNBr II was partially and CNBr III and IV were completely dissolved. Gel filtration on Biogel P-10 (200–400 mesh, 20 × 800 mm column) with 90% formic acid as solvent separated CNBr III–IV, CNBr III and CNBr IV. These three fragments could be sequenced completely in the liquid phase sequenator. They represent the C-terminus, which is identical with the bovine lipophilin C-terminal CNBr-fragments III and IV (Fig. 1). Their amino-acid compositions are summarized in the table.

CNBr II (5.6 kDa) was purified by repeated HPLC on coupled Si 60 and Si 100 columns. Automated Edman degradation over 19 cycles yielded the following sequence:

Tyr-Gly-Val-Leu-Pro-Trp-Asn-Ala-Phe-Pro-Gly-
-Lys-Val-Cys-Gly-Ser-Asn-Leu-Leu-

The amino-acid stoichiometry of CNBr II after a 24 h hydrolysis in 6M HCl is summarized in lane 2 of the table.

The above sequence is identical with the sequence Tyr^{206}-Leu^{224} of bovine PLP.

We noticed in our previous studies that cleavage at tryptophan residues with HBr/dimethyl sulfoxide is more appropriate for PLP, because a smaller Trp I (M_r 16 kDa) was formed for further degradation in addition to three medium sized polypeptides of Trp II M_r 2.0 kDa, Trp III M_r 5.3 kDa and Trp IV M_r 7 kDa.

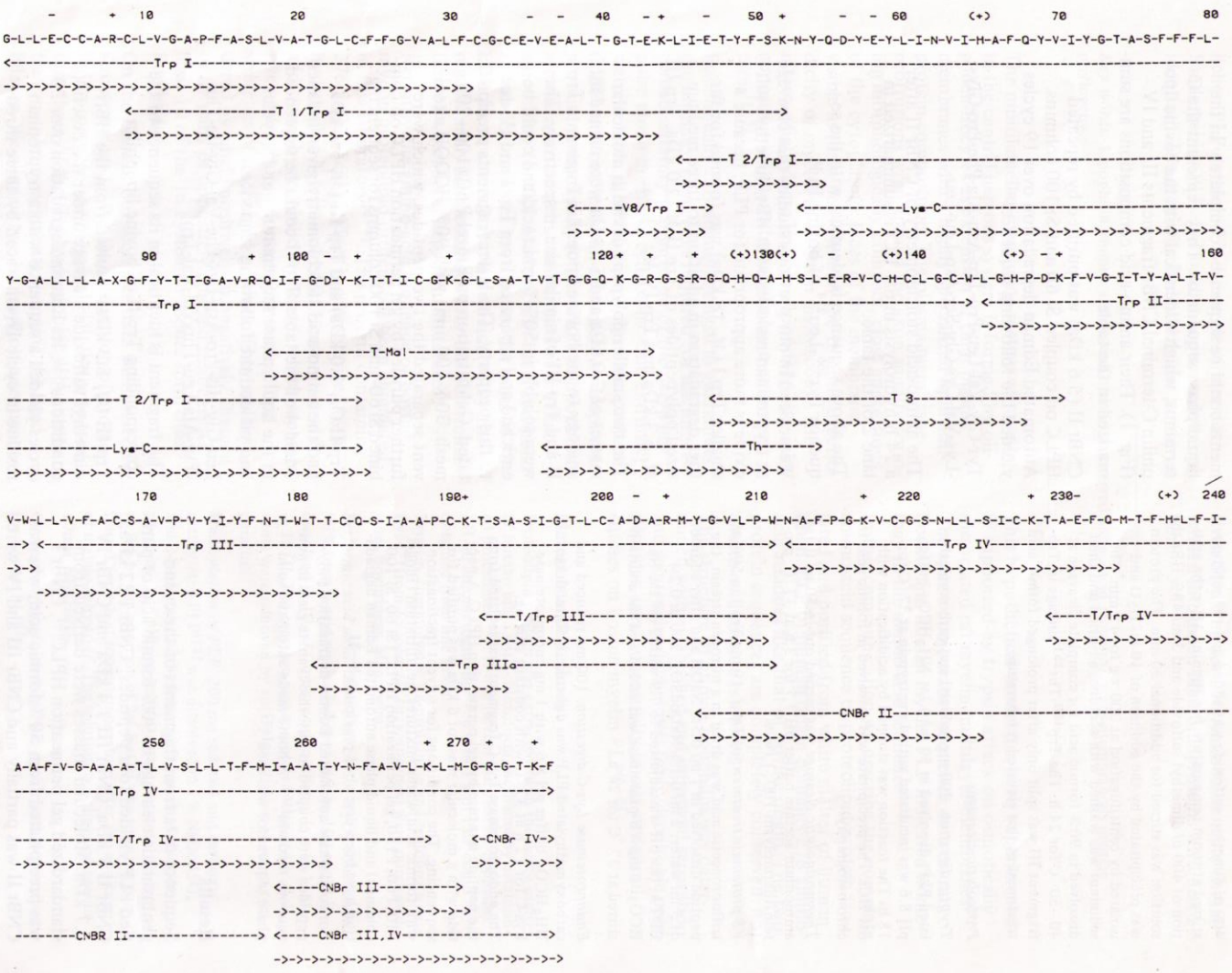
The two smaller polypeptides with molecular masses of 2.0 kDa and 5.3 kDa were named in analogy to the bovine lipophilin fragments Trp II and Trp III of similar size, respectively. They were soluble in 50% formic acid and could be enriched and separated from Trp I and IV due to this property. The latter fragments precipitated. Gel filtration on Biogel P 60 (100–200 mesh; 30 × 100 mm) with 90% HCOOH as solvent separated the two peptides, which were further purified by gel permeation HPLC on combined Si 60- and Si 100-columns for sequencing.

Trp II (M_r 2.0 kDa) and Trp III (M_r 5.3 kDa) had the amino-acid stoichiometry given in lanes 3 and 4 of the table. Trp II could be sequenced in the liquid phase sequenator over 18 cycles and yielded the following sequence:

Leu-Gly-His-Pro-Asp-Lys-Phe-Val-Gly-Ile-Thr-
-Tyr-Ala-Leu-Thr-Val-Val-Trp

The fragment is identical in its sequence to the corresponding Trp II of bovine lipophilin.

Trp III (M_r 5.3 kDa) resulted from the HBr/dimethyl sulfoxide cleavage under the described conditions[15]. Its Edman degradation over 20 cycles yielded a sequence with a tryptophan residue in cycle 18 embraced by three threonine residues, one toward the N- and two toward the C-terminus. We have earlier observed the resistance of the Trp-Thr linkage under the standard cleavage conditions[7,8].


G-L-L-E-C-C-A-R-C-L-V-G-A-P-F-A-S-L-V-A-T-G-L-C-F-F-G-V-A-L-F-C-G-C-E-V-E-A-L-T-G-T-E-K-L-I-E-T-Y-F-S-K-N-Y-Q-D-Y-E-Y-L-I-N-V-I-H-A-F-Q-Y-V-I-Y-G-T-A-S-F-F-L-
Trp I
T 1/Trp I
T 2/Trp I
V8/Trp I
Lys-C
Y-G-A-L-L-L-A-X-G-F-Y-T-T-G-A-V-R-Q-I-F-G-D-Y-K-T-T-I-C-G-K-G-L-S-A-T-V-T-G-G-Q- K-G-R-G-S-R-G-Q-H-Q-A-H-S-L-E-R-V-C-H-C-L-G-C-W-L-G-H-P-D-K-F-V-G-I-T-Y-A-L-T-V-
Trp II
T-Mal
T 2/Trp I
T 3
Lys-C
Th
V-W-L-L-V-F-A-C-S-A-V-P-V-Y-I-Y-F-N-T-W-T-T-C-Q-S-I-A-A-P-C-K-T-S-A-S-I-G-T-L-C- A-D-A-R-M-Y-G-V-L-P-W-N-A-F-P-G-K-V-C-G-S-N-L-L-S-I-C-K-T-A-E-F-Q-M-T-F-H-L-F-I-
Trp III
Trp IV
T/Trp III
T/Trp IV
Trp IIIa
CNBr II
A-A-F-V-G-A-A-A-T-L-V-S-L-L-T-F-M-I-A-A-T-Y-N-F-A-V-L-K-L-M-G-R-G-T-K-F
Trp IV
T/Trp IV
CNBr IV
CNBr III
CNBR II
CNBr III,IV

◁

Fig. 1. Amino-acid sequence of human brain myelin proteolipid (lipophilin).

Dotted lines under the sequence indicate the fragments obtained either by chemical (cyanogenbromide (CNBr) and *HBr/dimethyl sulfoxide (Trp)) or by proteolytic cleavage. Arrows indicate the sections sequenced* by Edman degradation. T1/Trp I and T2/Trp I = tryptic fragments of Trp I; V8/Trp I = splitting products of Trp I by *Staphylococcus aureus* V8 proteinase (EC 3.4.21.19); Lys-C = splitting product of Trp I by endoproteinase Lys-C (from *Lysobacter enzymogenes*); T-Mal = tryptic peptide from the malic anhydride-derivatized Trp I fragment; Th = thermolysin degradation product of the reduced and carboxymethylated proteolipid protein; T3 = a tryptic fragment derived from the whole proteolipid protein. + and − indicate charged side chains.

Table. Amino-acid analyses of fragments purified from chemical (cyanogen bromide and HBr/dimethyl sulfoxide) and proteolytic digestions of human white matter proteolipid protein.

	CNBr III, IV (258–276)		CNBr II (206–257)		Trp II (145–162)		Trp III (163–211)		
Asp	1.0	(1)	3.5	(2)	1.2	(1)	2.1	(2)	
Thr	1.0	(2)	2.8	(4)	1.6	(2)	3.1	(5)	
Ser	–	(–)	2.7	(3)	0.4	(–)	4.0	(4)	
Glu	–	(–)	2.3	(2)	0.7	(–)	0.8	(1)	
Pro[a]	–	(–)	–	(2)	–	(1)	–	(3)	
Gly	0.5	(2)	4.2	(4)	2.4	(2)	3.9	(2)	
Ala	3.4	(3)	5.5	(7)	1.2	(1)	5.5	(7)	
Val	1.3	(1)	2.9	(4)	3.2	(3)	5.0	(4)	
Met	–	(1)	0.3	(2)	–	(–)	0.2	(1)	
Ile	0.8	(1)	2.0	(2)	1.2	(1)	3.4	(3)	
Leu	2.1	(2)	5.2	(7)	0.6	(2)	4.0	(4)	
Tyr	0.6	(1)	0.8	(1)	–	(1)	–	(3)	
Phe	1.3	(2)	3.0	(6)	1.2	(1)	2.4	(2)	
Lys	0.9	(2)	2.1	(2)	0.8	(1)	1.0	(1)	
His	–	(–)	0.7	(1)	1.0	(1)	0.1	(–)	
Arg	0.3	(1)	1.0	(–)	–	(–)	1.1	(1)	
Trp[a]	–	(–)	–	(1)	–	(1)	–	(2)	
Cys[a]	–	(–)	–	(2)	–	(–)	–	(4)	
Total	13.2	(19)	39.0	(52)	15.5	(18)	36.6	(49)	

	T/Trp IV (229–268)		T/Trp III (192–211)		T-Mal (98–123)		T3 (127–150)		Lys-C (53–104)	
Asp	0.8	(1)	0.9	(1)	1.0	(1)	1.0	(1)	2.5	(4)
Thr	2.4	(5)	0.8	(2)	3.6	(4)	–	(–)	2.4	(3)
Ser	–	(1)	1.1	(2)	2.1	(1)	0.7	(1)	0.7	(1)
Glu	0.7	(2)	0.1	(–)	1.8	(2)	2.9	(3)	2.8	(4)
Pro[a]	–	(–)	–	(1)	–	(–)	–	(1)	–	(–)
Gly	1.5	(1)	0.9	(2)	5.0	(6)	3.5	(3)	5.1	(5)
Ala	7.6	(9)	2.0	(3)	1.1	(1)	1.0	(1)	4.0	(5)
Val	2.8	(3)	1.0	(1)	1.2	(1)	1.0	(1)	2.9	(3)
Met	–	(2)	–	(1)	–	(–)	–	(–)	–	(–)
Ile	1.8	(2)	1.1	(1)	2.0	(2)	0.4	(–)	2.9	(4)
Leu	4.7	(5)	1.9	(2)	1.1	(1)	2.4	(3)	5.1	(5)
Tyr	–	(1)	–	(1)	1.1	(1)	–	(–)	6.1	(8)
Phe	4.0	(6)	0.1	(–)	1.9	(1)	0.1	(–)	4.8	(6)
Lys	1.0	(1)	0.2	(–)	2.8	(3)	0.4	(1)	0.2	(1)
His	0.1	(1)	–	(–)	–	(–)	3.8	(4)	0.4	(1)
Arg	0.2	(–)	0.7	(1)	0.5	(1)	0.9	(1)	0.8	(1)
Trp[a]	–	(–)	–	(1)	–	(–)	–	(1)	–	(–)
Cys[a]	–	(–)	–	(1)	–	(1)	–	(3)	–	(–)
Total	27.7	(40)	10.8	(20)	25.2	(26)	18.1	(24)	40.7	(51)

[a] Proline is not detected by *o*-phthaldialdehyde, Trp and Cys deteriorate upon acid hydrolysis and are not detected. Theoretical values in parenthesis. They were confirmed by automated Edman degradation.

However, a more drastic and repeated cleavage released a Trp-fragment, IIIa, from Trp III. 14 cycles in automated Edman degradation revealed the following N-terminus of Trp IIIa:

Thr-Thr-Cys-Gln-Ser-Ile-Ala-Ala-Pro-Cys-Lys-
-Thr-Ser-Ala-

The complete sequence of Trp IIIa resulted from the sequencing of a trypsin fragment (T/Trp III) isolated from a PLP tryptic digest of M_r 2.1 kDa over 19 cycles:

Thr-Ser-Ala-Ser-Ile-Gly-Thr-Leu-Cys-Ala-Asp-
-Ala-Arg-Met-Tyr-Gly-Val-Leu-Pro

As indicated in Fig. 1 fragment T/Trp III overlaps with the sequence obtained from the N-terminal end of Trp III.

The Trp IV fragment was enriched in the sediment insoluble in 50% formic acid together with Trp I from which it was purified by gel-permeation HPLC (Si 100–60) and further degraded with trypsin. The amino-acid analysis indicated the presence of one lysine residue. The automated Edman degradation allowed reliable PTH-identification over only 22 cycles with the lysine residue in position 17:

Asn-Ala-Phe-Pro-Gly-Lys-Val-Cys-Gly-Ser-Asn-
-Leu-Leu-Ser-Ile-Cys-Lys-Thr-Ala-Glu-Phe-Gln-

The trypsin fragment of M_r 4.3 kDa (T/Trp IV) was coupled in a non-aqueous medium to solid-phase-controlled pore glass support (PITC-CP-glass) and could be sequenced over 33 cycles. This sequence was unravelled. It overlapped with the C-terminus of Trp IV and with the N-terminus of CNBr III at its C-terminus (bold-type sequences):

Thr-Ala-Glu-Phe-Gln-Met-Thr-Phe-His-Leu-Phe-
-Ile-Ala-Ala-Phe-Val-Gly-Ala-Ala-Ala-Thr-Leu-
-Val-Ser-Leu-Leu-Thr-Phe-**Met-Ile-Ala-Ala-Thr-**

Trp I (M_r appr. 16–17 kDa) was obtained in homogenous form after repeated HPLC of the fraction insoluble in 50% HCOOH. The fragment was oxidized with performic acid. The amino-terminal 31 amino-acid residues were identical with the N-terminus of bovine lipophilin. Trypsin treatment of the performic acid-treated Trp I fragment in its unmodified form yielded a 4.6-kDa peptide which was purified by HPLC (Si 100–60) and sequenced over 35 cycles. Tryptic cleavage had occurred at Arg^8 of Trp I and 23 residues of the T1/Trp I fragment overlapped with the N-terminus of Trp I. Another trypsin fragment (T2/Trp I) was sequenced over 10 cycles, two of which overlapped with the C-terminus of a V8-proteinase fragment of Trp I. Based on our experience with bovine lipophilin we applied the endoproteinase Lys-C to Trp I of human myelin with the same success. The endoproteinase Lys-C fragment could be sequenced over 52 cycles. Two residues at the N-terminus overlapped with the C-terminus of T2/Trp I and seven C-terminal residues with the tryptic peptide from the maleic anhydride-derivatized Trp I fragment, the 26 residues sequence of which was established by liquid-phase Edman degradation.

Trp I was also hydrolysed with *Staph. aureus* V8 proteinase, which released a 1-kDa peptide. After HPLC (Si 100–60) purification it yielded the sequence:

Ala-Leu-Thr-Gly-Thr-Glu-Lys-Leu-Ile-

which closed the gap betwen T1/Trp I and T2/Trp I. The sequence data of Trp I still indicated a discontinuity of a 2.5 kDa sequence between the C-terminus of T-Mal/Trp I and the N-terminus of Trp II.

Since all sequences of human lipophilin had so far proved to be identical with those of the bovine proteolipid protein, we assumed that the gap was within a long hydrophilic sequence. Treatment of reduced and carboxymethylated lipophilin with thermolysin released a 2.1-kDa peptide (Th) the sequenced 16 N-terminal residues of which overlapped with the C-terminus of T-Mal/Trp I and with the N-terminus of a tryptic fragment T3 (2.6 kDa). The peptide T3 was isolated from the tryptic peptide mixture of lipophilin (see above), it was soluble in 10% formic acid and purified on Biogel P30 (100–200 mesh) and by HPLC (Si 100–60). From the automated Edman degradation over 21 cycles the following sequence was derived:

Gly-Gln-His-Gln-Ala-His-Ser-Leu-Glu-Arg-Val-
-Cys-His-Cys-Leu-Gly-Cys-Trp-**Leu-Gly-His-**

The array of fragments purified from chemical and enzymatic cleavage procedures and sequenced either by liquid or solid phase automated Edman degradation led to the sequence of the complete amino-acid sequence of human lipophilin, the main integral membrane protein of central nervous system myelin. It consists of 276 amino acids with an exact molecular mass of 29 728 Da.

The human PLP sequence is absolutely identical with the bovine PLP.

Discussion

The amino-acid sequence analysis of human proteolipid apoprotein (lipophilin) described in this paper revealed the suprising result that this integral, extremely hydrophobic membrane

protein is identical with the bovine lipophilin, the primary structure of which we have recently described[6–9]. In this sequence position 254 is leucine instead of valine and X in position 36 has been identified as valine. Lipophilin of both species does not show a single homologous or non-homologous amino-acid exchange, provided that residue 88, unidentified in both species, proves to be identical. It remains to be analysed whether the same conservative structure is preserved on the DNA level of the two species.

The structure of human lipophilin given in Fig. 1 reveals that the sequence with 276 amino-acid residues is divided into domains of distinctly different polarities. Five hydrophobic segments form clusters of 26, 32, 40, 12, and 30 hydrophobic amino acids. These domains with uncharged, hydrophobic amino acids are recognized when the hydropathy of lipophilin is plotted using relative hydrophobicity indices (HI)[18,19] (Fig. 2). We have proposed a model in which lipophilin is oriented as a polytopic integral membrane protein[9] (Fig. 3).

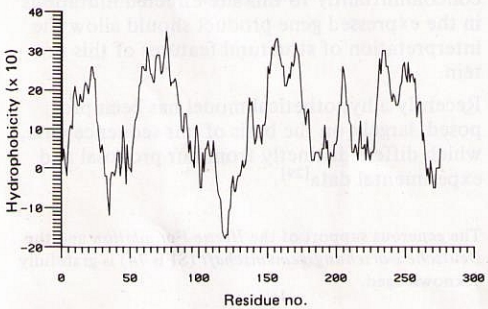


Fig. 2. Hydropathy plot of human brain myelin proteolipid protein (lipophilin).

The ordinate resembles the relative hydrophobicity indices (x 10) according to ref.[18].

The five hydrophobic domains are embedded in the lipid bilayer. Whereas about 21 amino-acid residues in α-helical arrangement are required to span a 3 nm bilayer[20], the thickness of the myelin membrane bilayer is about 4.6 nm according to electron microscopic and X-ray diffraction data[21,22]. A sequence of approximately 30 amino-acid residues in α-helical order is required for a transmembrane segment.

Our model, first based on theoretical reasoning and later supported by proteolytic cleavage experiments on dissociated myelin membranes[9] places the N-terminus on the external face, the

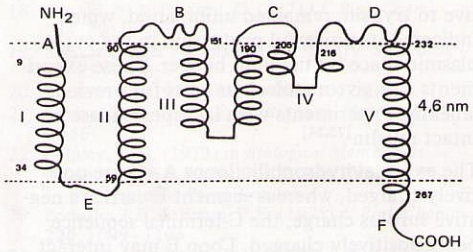


Fig. 3. Proposed assembly of proteolipid protein (lipophilin) in the lipid bilayer of the myelin membrane.

Segments I, II and V are trans-membranous helices, III and IV cis-membranous segments. The N-terminus A and loops B, C and D are hydrophilic and oriented toward the external surface of the membrane, loop E and the C-terminus are located on the cytosolic side of the lipid bilayer. The width of the lipid bilayer is assumed to be 4.6 nm[20,21].

C-terminus on the cytoplasmic side of the membrane (Fig. 3). The hydrophobic domains I (Cys^{9}-Cys^{34}), II (Tyr^{59}-Phe^{90}) and V (Phe^{232} or Leu^{238}-Leu^{267}) are transmembranous segments, whereas domains III (Phe^{151}-Cys^{190}) and IV (Met^{205}-Gly^{216}) are hydrophobic cis-membranous segments. The hydrophobic segment IV (12 residues) is to short and domain III to large (40 residues) to span the bilayer. Furthermore Pro^{172} in domain III and Pro^{210} of domain IV may induce interior β-turns in the lipid bilayer. In addition we have established a disulfide linkage in bovine lipophilin connecting the hydrophilic loop between Gly^{216} and Glu^{231} with the N-terminal hydrophilic sequence Gly^{1}-Arg^{8} each of which embraces two cysteine residues[10]. As the participating cysteine residues must be on the same side of the membrane, which is the external side in our model, this implies that domain III enters the lipid phase, makes a β-turn and reemerges on the same (external) side.

Lipophilin contains 16 cysteine residues, 12 of which are in the hydrophilic intervening loops A, B, C and D located outside of the membrane bilayer, whereas four cysteine residues are located in the transmembrane segments I and III, but none on the cytoplasmic side.

The proposed orientation of the polypeptide chain with its hydrophilic and hydrophobic trans- and cis-membranous segments was supported by treatment of the hyposmotically dissociated multi-membrane layer of myelin with trypsin by which loops B and C were cleaved resulting in three polypeptides, which were separated and determined in their structure[9]. Basic myelin protein, which is extremely sensi-

tive to trypsin, remained unimpaired, which indicated its powerful protection in the cytoplasmic space by the lipid bilayer. These experiments also give a molecular basis for previous labelling experiments with lactoperoxidase on intact myelin[23,24].

The external hydrophilic loops A–D are positively charged, whereas segment E carries a negative surplus charge, the C-terminal sequence being positively charged. Loop E may interact with MBP, whereas the positively charged loops may contribute to the tight packing of the myelin membranes by electrostatic interaction with the acidic polar head group of the bilayer lipids (phosphatidylinositol, phosphatidylserine and sulfatides).

Although the exact localisation of the disulfide linkages has still to be clarified, the previously indicated disulfide linkage between half cystines at the N-terminus and the hydrophilic loop D toward the C-terminus suggests that a contraction of the hydrophilic loops by disulfide bonds between loops leads to a clustering of transmembranous segments. We favour the α-helical over the β-sheet structures of the hydrophobic domains, which is suggestive for thermodynamic reasons though not proven, because intrahelical hydrogen bonding is possible only in an α-helix but not in a β-sheet structure, exposing either hydrophobic or polar neutral side chains of the hydrophobic segments for interactions[25].

The hydrophobic sequences of the transmembranous domains I, II and IV and of the cismembranous domain III are rich in aromatic side chains, which in molecular models show amphipathic sidedness and show stacking of the aromatic side chains, and are also prone to intercalation of aromatic side chains of neighbouring helices. Likewise the neutral polar side chains are oriented to a large extent towards one side and require an interaction with a similarly oriented helix of the same or another PLP molecule within the lipid bilayer in order to mask the polarity of these side chains. Future chemical experiments will demonstrate whether lipophilin molecules aggregate within the membrane as suggested recently[26,27] and whether this integral membrane protein forms the basis of the radial components visible in electron microscopy[28]. Two further structural features warrant a discussion with regard to the myelin multilayer membrane structure.

In our elucidation of bovine proteolipid protein we discovered that threonine-198 within the hydrophilic segment C on the external side of the membrane was acylated with a long chain fatty acid.

This hydrophilic loop is followed by a 12-residue-long cis-hydrophobic domain IV with a β-turn. With our present techniques we cannot ascertain whether the fatty acid residue[5] and the hydrophobic cis-domain IV immerge into the outer layer of the same or the opposite membrane bilayer. These two features could contribute hydrophobic in addition to ionic interactions to the tight packing of the myelin membranes wrapped around the axon. The amino-acid sequences of the hydrophilic domains, four of which according to our model are situated on the external side of the membrane, as well as those of the hydrophobic membrane-integrated domains are completely conserved in man and cattle. The very limited sequence data of other integral membrane proteins do not allow a comparison with the results of our structural studies on myelin integral proteolipid protein of bovine and human brain, which revealed identical amino-acid sequences. The significance of the conserved protein structure with regard to the function can only be deduced if a three-dimensional structure of lipophilin within the lipid bilayer and on its surface becomes available. Prior or concommittantly to this site-directed mutations in the expressed gene product should allow the interpretation of structural features of this protein.

Recently a hypothetical model has been proposed, largely on the basis of our sequence data, which differs distinctly from our proposal and experimental data[29].

The generous support of the *Hertie-Foundation* and the *Deutsche Forschungsgemeinschaft* (SFB 74) is gratefully acknowledged.

Literature

1 Norton, W.T. (1981) in *Basic Neurochemistry* (Siegel, G.J., Albers, R.W., Agranoff, B.W. & Katzman, R., eds.) 3rd edn., pp. 63–92, Little and Brown Comp., Boston.
2 Rumsby, M.G. (1978) *Biochem. Soc. Trans.* **6**, 448–462.
3 Carnegie, R.R. & Moore, W.J. (1981) in *Proteins of the Nervous System* (Bradshaw, R.A. & Schneider, D.M., eds.) 2nd edn., pp. 119–143, Raven Press, New York.
4 Oshiro, Y. & Eylar, E.H. (1970) *Arch. Biochem. Biophys.* **138**, 606–613.
5 Eylar, E.H., Brostoff, S., Hashim, G., Caccam, J. & Burnett, J.P. (1971) *J. Biol. Chem.* **246**, 5770–5784.
6 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 855–864.
7 Stoffel, W., Schröder, W., Hillen, H. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 1117–1131.

8 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 1397–1407.
9 Stoffel, W., Hillen, H. & Giersiefen, H. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5012–5016.
10 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1983) *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 1455–1466.
11 Lees, M.B., Sakura, J.D., Sapirstein, V.S. & Curatolo, W. (1979) *Biochim. Biophys. Acta* **559**, 209–230.
12 Folch-Pi, J. & Lees, M.B. (1951) *J. Biol. Chem.* **191**, 807–817.
13 Butler, P.J.G., Harris, J.I., Hartley, B.S. & Leberman, R. (1969) *Biochem. J.* **112**, 679–689.
14 Gross, E. (1967) *Methods Enzymol.* **25**, 238–255.
15 Wachter, E. & Werhan, R. (1979) *Anal. Biochem.* **97**, 56–64.
16 Machleidt, W., Hofer, H. & Wachter, E. (1975) in *Solid-Phase Methods in Protein Sequence Analysis* (Laursen, R.A. ed.) pp. 17–30, Pierce Chemical Co., Rockford, Ill.
17 Machleidt, W. & Hofer, H. (1980) in *Methods in Peptide and Protein Sequence Analysis* (Birr, Chr. ed.) pp. 35–47, Elsevier/North-Holland, Amsterdam.
18 Nozaki, Y. & Tanford, C. (1971) *J. Biol. Chem.* **246**, 2211–2217.
19 Segrest, J.P. & Feldman, R.J. (1974) *J. Mol. Biol.* **87**, 853–858.
20 Tanford, C. (1978) *Science* **200**, 1012–1018.
21 Levine, Y.K. (1975) *Prog. Biophys. Mol. Biol.* **29**, 3–56.
22 Shipley, G.G. (1973) in *Biological Membranes* (Chapman, D. & Wallach, D.F., eds.) Vol. 2, pp. 1–89, Academic Press, London.
23 Poduslo, J.F. & Braun, P.E. (1975) *J. Biol. Chem.* **220**, 1099–1105.
24 Golds, E.E. & Braun, P.E. (1978) *J. Biol. Chem.* **253**, 8162–8170.
25 Gray, T.M. & Matthews, B.W. (1984) *J. Mol. Biol.* **175**, 75–81.
26 Harris, R. & Fridlay, J.B.C. (1983) *Biochim. Biophys. Acta* **732**, 75–82.
27 Smith, R., Cook, J. & Dickens, P.A. (1984) *J. Neurochem.* **42**, 306–313.
28 Rumsby, M.G. & Crang, A.J. (1977) *Cell Surf. Rev.* **4**, 247–362.
29 Laursen, R.A., Samiullah, M. & Lees, M.B. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2912–2916.

Prof. Dr. Dr. Wilhelm Stoffel, Helmut Giersiefen, Dipl. Chem., Dr. Heinz Hillen, Dr. Werner Schroeder and Dr. Budi Tunggal, Institut für Physiologische Chemie der Universität zu Köln, Joseph-Stelzmann-Str. 52, D-5000 Köln 41.