Hoppe-Seyler's Z. Physiol. Chem.
Bd. 364, S. 1455–1466, Oktober 1983

The Primary Structure of Bovine Brain Myelin Lipophilin (Proteolipid Apoprotein)

Wilhelm STOFFEL, Heinz HILLEN, Werner SCHRÖDER and Rainer DEUTZMANN

Institut für Physiologische Chemie der Universität zu Köln

(Received 9 May/25 August 1983)

Summary: The amino-acid sequence of bovine myelin lipophilin (proteolipid apoprotein, Folch-protein) has been completed. Lipophilin is a 276 amino acid residues containing, extremely hydrophobic membrane protein with molecular mass 30 000 Da. The sequence determination was based on automated Edman degradation of four tryptophan and four cyanogen bromide fragments and of proteolytic peptides of complete lipophilin as well as the fragments obtained by chemical cleavage.

Four additional sequences were determined which led to the completion of the primary structure. Lipophilin is esterified at threonine-198 by long chain fatty acids (palmitic, stearic and oleic acid). The attachment site has been established at the same threonine residue in three different peptides isolated from thermolysinolytic, papainolytic and chymotrypsinolytic hydrolysates. This threonine residue is part of a hydrophilic segment of lipophilin.

The covalent fatty acyl bond is being discussed together with important structural and functional properties of this membrane protein which can be derived from sequence information.

New separation and purification methods of hydrophobic and hydrophilic polypeptides for this sequence determination (fractional solubilization, silica gel exclusion, high-performance liquid chromatography) had to be elaborated as indispensable tools. They are generally applicable to the structural analysis of hydrophobic membrane proteins.

Four long (26, 29, 40 and 36 residues) and one medium long (12 residues) hydrophobic segments are separated by four predominantly positively and one negatively charged hydrophilic segments. On the basis of structural data a model for the membrane integration of lipophilin is proposed.

Die Primärstruktur des Lipophilins (Proteolipid-Apoprotein) aus Rindergehirnmyelin

Zusammenfassung: Die Aminosäure-Sequenz des Myelin-Lipophilins (Proteolipid-Apoprotein, Folch-Protein) des Rindergehirns wurde bestimmt. Die Sequenzierung dieses äußerst hydrophoben, 276 Aminosäurereste enthaltenden Membranproteins der Molekularmasse 30 000 Da,

Enzymes:
Chymotrypsin (EC 3.4.21.1); papain (EC 3.4.22.2); thermolysin (EC 3.4.24.4); trypsin (EC 3.4.21.4); Staphylococcal serine proteinase (*Staph. aureus* strain V8 proteinase) (EC 3.4.21.19); endoproteinase Lys-C (EC 3.4.21.-).
Abbreviation:
BNPS-skatol = 3-bromo-3-methyl-2-(2-nitrophenylsulfenyl)-3*H*-indole.

basierte auf dem automatischen Edman-Abbau von vier Tryptophan- und vier Bromcyanbruchstücken sowie den proteolytischen Spaltstücken des Lipophilins. Vier weitere Teilsequenzen wurden bestimmt, die zur Vervollständigung der Primärstruktur führten.

Lipophilin enthält eine langkettige Fettsäure, esterartig gebunden an Threonin-198. Diese Bindungsstelle wurde in drei verschiedenen Peptiden, die durch Thermolysin-, Papain- und Chymotrypsinspaltung gewonnen wurden, bestimmt. Die Bindungsstelle befindet sich an einem Threoninrest eines hydrophilen Segmentes des Lipophilins.

Die kovalente Fettsäurebindung wird im Zusammenhang mit den durch die Sequenzinformationen erhaltenen wichtigen Eigenschaften für die Struktur und Funktion dieses Membranproteins diskutiert.

Das Lipophilin enthält vier lange, hydrophobe Segmente (von 26, 29, 40 und 36 Aminosäuren) und einen kürzeren, hydrophoben Abschnitt (12 Aminosäuren). Diese hydrophoben Bereiche sind durch hydrophile Aminosäuresequenzen getrennt. Auf Grund dieser Befunde wird ein Modell für die Anordnung des Lipophilin-Moleküls in der Membran vorgeschlagen.

Key words: Membrane protein sequencing; lipophilin (proteolipid apoprotein), purification and Edman degradation of hydrophobic polypeptides; sites of proteolytic cleavage; asymmetric assembly of lipophilin into membrane.

The two major structural proteins of the myelin membrane in central nervous system are proteolipid apoprotein, also named lipophilin[1], and basic myelin protein. Lipophilin comprises more than 50% of the total myelin proteins.

Whereas the primary structure of basic myelin protein has been elucidated[2], lipophilin, first described in 1951[3], resisted all attempts to elucidate its primary structure because of several unusual properties of this chloroform-methanol soluble membrane protein. Many aspects of myelin proteolipid have been compiled[4,5]. Due to the strong interactions with lipids and with itself and to the conformational flexibility of lipophilin, depending on the polarity of the solvent, a strong tendency of aggregation can be observed and has led to reports suggesting an oligomeric structure with subunits of 12.5 kDa[6] and 5 kDa, respectively[7]. Most studies, however, suggested molecular masses of 23.5, 25, 28 and 30 kDa[8–11].

The hydrophobicity of lipophilin is further enhanced by covalently bound long-chain fatty acids (1–2% corresponding to 1–2 mol fatty acid/mol lipophilin)[12–14].

Chemical studies on lipophilin of rat brain were initiated in 1974 by Edman degradation of the N-terminus, which proved to be homologous in lipophilin of human[1,16] and bovine white matter[1,17,18].

Cyanogen bromide cleavage releases four fragments, two of which in automated Edman degradation yielded 13 and 6 residues long sequences, the latter with the C-terminal phenylalanine residue[9]. The alignment could be established by sequencing the corresponding 19 residues long fragment obtained by incomplete BrCN-cleavage. A third fragment could only be degraded partially.

Small tryptic peptides of lipophilin have been sequenced[19] and recently the sequence of a tryptic fragment (4 kDa), released from the C-terminus, has been established[20].

In addition to the cyanogen bromide cleavage, we applied the chemical cleavage with HBr/dimethyl sulfoxide[21] and BNPS-skatol[22]. Gel permeation and gel permeation high-performance liquid chromatographic procedures were elaborated in this laboratory for the separation of hydrophobic and hydrophilic peptides, which allowed us to separate the four cyanogen bromide and four HBr/dimethyl sulfoxide fragments and purify them for automated Edman degradation. In addition, gas-liquid chromatographic/mass spectrometric analysis of the derivatized peptides of the partial acid hydrolysate was utilized.

These studies unravelled the sequence of cyanogen bromide fragments II, III and IV at the C-terminus (71 amino acid residues) which comprise the C-terminal HBr/dimethyl sulfoxide fragment IV (Trp IV, 65 amino acid residues)[23]. Cleavage

with BNPS-skatol was also applied to lipophilin of rat brain myelin[24]. The partial sequence of tryptophan fragment I and the complete fragments II, III and IV were determined by automated Edman degradation[25].

Three large and several smaller tryptic and thermolysinolytic fragments were isolated and purified for sequence determination. They form partial structures of the 16 kDa large tryptophan fragment I or the 23 kDa large cyanogen bromide fragment I[26], the lengths of which are now firmly established.

Our sequence data left two gaps in the amino-acid sequence of cyanogen bromide fragment I and Trp fragment I, respectively. This paper reports the amino-acid sequence of bovine white matter lipophilin. It consists of 276 amino-acid residues with a molecular mass of 30 kDa.

Threonine-198 is esterified with long chain fatty acids (palmitic, stearic and oleic acid). Lipophilin contains 16 cysteine residues some of which are involved in disulfide bridges. A detailed summary of the strategy and experimental results which led to this extremely hydrophobic membrane protein will be given in a forthcoming publication.* Five strongly hydrophobic sequences, 26, 29, 40, 12 and 36 amino-acid residues long, which account for 50% of the amino-acid residues of lipophilin, are separated by hydrophilic sequences. Based on the cleavage sites of several proteolytic enzymes, on disulfide bonds between proteolytic fragments and the necessity of embedding the long hydrophobic sequence into the lipid bilayer of myelin, we propose a model for the assembly of lipophilin in the brain white matter myelin membrane.

Materials and Methods

Materials

Thermolysin was obtained from Serva (D-6900 Heidelberg), trypsin treated with tosyl-L-phenylalanyl-chloromethane from E. Merck (D-6100 Darmstadt), *Staph. aureus* strain V8 proteinase from Miles (D-6000 Frankfurt), papain and chymotrypsin from Sigma Chemie (D-8028 Taufkirchen), endoproteinase Lys-C from Boehringer (D-6800 Mannheim).

* Stoffel, W., Hillen, H., Schröder, W. and Deutzmann, R., this journal, in preparation.

Isolation and fragmentation of lipophilin

Bovine lipophilin was isolated as described[3]. Reductive carboxymethylation[27] and performic acid oxidation[28], cyanogen bromide cleavage at methionine[29] and HBr/dimethyl sulfoxide[21] or BNPS-skatol cleavage at tryptophan residues[22] were modified and adapted to this hydrophobic protein as described before[23,25,26]. For blocking α- and ε-amino groups maleoylation was performed at pH 9–11[29].

Proteolytic cleavage: Conditions for the proteolytic cleavage of lipophilin by trypsin and thermolysin of carboxymethylated lipophilin were those reported before[25]. Cleavage by the staphylococcal proteinase was performed in 0.1M ammonium hydrogen carbonate buffer pH 7.8, 2mM EDTA at 37 °C for 24 h at a substrate to enzyme ratio of 30:1[30].

Papain cleavage: Lipophilin (100 mg) was submitted to exhaustive reductive carboxymethylation as described earlier[25] and then cleaved with 5 mg of papain for 1 h at 37 °C. The reaction mixture was acidified with 98% formic acid to pH 2. The large peptides precipitated. They were extracted with 75% formic acid and chromatographed on a Bio-Gel P 100 column (2.5 × 120 cm), solvent: 90% formic acid. Fractions were analysed by high-performance liquid chromatography. Those eluting with identical retention times were pooled and rechromatographed (Si 100 and Si 60, 30 × 0.9 cm columns) for automated Edman degradation.

Chymotrypsin cleavage: 200 mg maleoylated and reductively carboxymethylated lipophilin was treated with 4 mg chymotrypsin for 24 h at room temperature in a total volume of 60 ml. The reaction was stopped by lyophilisation. After chromatography on Bio-Gel P 100 (conditions see above), the fractions with identical retention times were isolated and purified by repeated high-performance liquid chromatography.

Endoproteinase Lys-C cleavage: 40 mg reductively carboxymethylated lipophilin was treated with 3 U endoproteinase Lys-C for 18 h at 37 °C in 0.1M NH_4HCO_3 buffer, pH 8.6.

Separation, purification and sequencing of hydrophobic and hydrophilic peptides

Large peptides were precipitated by acidifying with 98% formic acid, centrifuged, the precipitate extracted with 75% formic acid and chromatographed on Bio-Gel P 30, P 60, P 100 or P 150 columns (2 × 120 cm) (solvent: 90% formic acid[23,25]). Fractions were analysed and purified by high-performance liquid chromatography as described[25]. For automated Edman degradation[31] a Beckman sequenator, model 890 C, was used with the 0.2M Quadrol program. Approximately 50 nmol was used for automated Edman degradation.

High-performance liquid chromatography of phenylthiohydantoin derivatives of amino acids was carried out on mixed bed nitrile and phenyl derivatized phases (1.2 g Nucleosil CN 5 μm and 2 g Nucleosil phenyl 7 μm) in stainless steel columns with the following buffer system: Buffer A: 2.9mM sodium acetate pH 5.0, Buffer B: 50% buffer A, 50% acetonitrile. Gradient: 25% B 1 min; 25% B to 72.5% B 10 min; 72.5% B 14 min; 72.5% B to 25% B 1 min; hold at 25% B for 10 min, temperature 32 °C.

Isolation and analysis of fatty acid, esterified to lipophilin

Phenylthiohydantoin fractions of serine and threonine residues positioned by automated Edman degradation of HBr/dimethyl sulfoxide fragments in previous sequencing were subsequently isolated from proteolytically released peptides of carboxymethylated lipophilin. The phenylthiohydantoin fractions, in which the derivatives of serine or threonine were identified, were saponified in 1M NaOH for 30 min at 80 °C. Fatty acids were extracted with hexane after acidification with 2M HCl and esterified in anhydrous methanolic HCl (5%) for 30 min at 80 °C. Methyl esters were separated and identified by gas-liquid chromatography (2 m, 2.5 mm diameter, 7.5% polyethylenglycol succinate at 140 °C).

Results

The amino-acid composition of bovine proteolipid protein after total hydrolysis is shown in Tab. 1. In our recent publications on its sequence[23,25,26] we described the chemical fragmentation by cyanogen bromide and HBr/dimethyl sulfoxide and the isolation of several fragments liberated by proteolysis of lipophilin. Four cyanogen bromide fragments and four HBr/dimethyl sulfoxide fragments (Trp fragments) were released, separated and purified for automated Edman degradation by a novel combination of methods: larger fragments, particularly those of hydrophobic nature, precipitated when the water content in formic acid exceeded 40%. Both the larger and smaller peptides were partially separated on Bio-Gel of respective pore sizes and further purified by repeated gel permeation

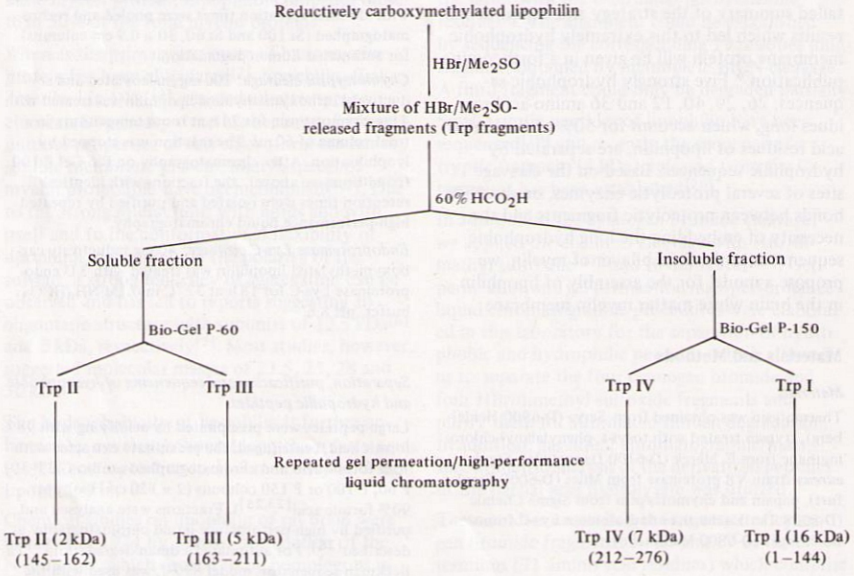

Fig. 1. Flow diagram of isolation of chemically released polypeptide fragments of lipophilin. In principle cyanogen bromide fragments were isolated by the same sequence of separation procedure.

Table 1. Amino acid composition of lipophilin. Conditions of hydrolysis: 6M HCl, 24 h at 110 °C in ampoules sealed under nitrogen.

(a) Calculated from sequence data; (b) amino acid analysis, based on calculation for 31 alanine residues.

Amino acid	a	b
Ala	31	31.0
Arg	7	7.1
Asx	10	11.5
Cys	16	9.7*
Glx	16	16.2
Gly	28	26.6
His	6	5.8
Ile	13	12.8
Leu	29	29.0
Lys	11	12.0
Met	4	3.2
Phe	22	21.8
Pro	6	7.8
Ser	13	12.8
Thr	24	21.2
Trp	4	–
Tyr	14	13.4
Val	21	18.7

* Determined as *S*-carboxymethylcysteine.

high-performance liquid chromatography on silicic acid with 90% formic acid as solvent. The flow sheet, Fig. 1, outlines as an example, the different steps leading to the isolation and purification of tryptophan fragments of lipophilin. Proteolytic cleavage reactions on lipophilin had indicated that the specific sites were localized predominantly in hydrophilic sequences[25]. The combination of the chemical and enzymatic cleavage reactions produced ample overlapping fragments to yield the unambiguous arrangement of the four fragments released by BNPS-skatol or HBr/dimethyl sulfoxide and cyanogen bromide, although the amino-acid sequence of the large N-terminal Trp-fragment I* (16 kDa) was incomplete. Further proteolytic digestions resulted in the completion of the sequence of TrpI. Exhaustive cleavage of drastically reduced and carboxymethylated lipophilin yielded a 3.6-kDa fragment T1, which was sequenced over 30 cycles:

T1:

$$\overset{10}{\text{Cys}}\text{-Leu-Val-Gly-Ala-Pro-Phe-Ala-Ser-Leu-Val-}\overset{20}{\text{Ala}}\text{-Thr-}$$
$$\text{-Gly-Leu-Cys-Phe-Phe-Gly-Val-Ala-}\overset{30}{\text{Leu}}\text{-Phe-Cys-Gly-Cys-}$$
$$\text{-Glu-(Val)-}\overset{38}{\text{Glu}}\text{-Ala-}$$

After exhaustive trypsin cleavage of maleoylated, reductively carboxymethylated lipophilin a peptide T-Mal-1 of 2.5 kDa molecular mass was isolated by the described procedure and sequenced over 22 cycles:

T-Mal-1:

$$\overset{98}{\text{Gln}}\text{-Ile-}\overset{100}{\text{Phe}}\text{-Gly-Asp-Tyr-Lys-Thr-Thr-Ile-Cys-Gly-}\overset{110}{\text{Lys}}\text{-}$$
$$\text{-Gly-Leu-Ser-Ala-Thr-Val-Thr-Gly-}\overset{119}{\text{Gly}}\text{-}$$

Endoproteinase Lys-C cleavage of completely reduced lipophilin yielded a peptide Lys-C-1 (5.6 kDa), which was sequenced over 52 cycles:

Table 2. Stoichiometry of amino acids of the enzymatic fragments.

The amino acids were detected fluorometrically with *o*-phthalaldehyde. Cysteine was determined as *S*-carboxymethylcysteine. Conditions of hydrolysis: 6M HCl, 24 h at 110 °C, in ampoules sealed under nitrogen. Values obtained from the sequence are given in brackets.

	Fragment			
Fragment	T1	T1-TrpI	V8-TrpI	Lys-C-1
Position	9–44	45–97	38–47	53–104
Mol. mass [kDa]	3.6	5.4	1.1	5.6
Asp	–	3.2 (3)	–	4.2 (4)
Thr	2.9 (3)	3.8 (4)	1.9 (2)	2.8 (3)
Ser	0.7 (1)	1.8 (2)	–	0.8 (1)
Glu	3.3 (3)	3.9 (4)	2.1 (2)	5.9 (4)
Gly	4.2 (5)	4.1 (4)	0.9 (1)	4.6 (5)
Ala	4.2 (5)	4.6 (5)	1.1 (1)	4.7 (5)
Val	3.4 (4)	2.3 (3)	–	3.3 (3)
Met	–	–	–	–
Ile	–	4.0 (4)	1.0 (1)	4.0 (4)
Leu	6.1 (5)	5.2 (6)	2.2 (2)	5.1 (5)
Tyr	–	8.1 (8)	–	8.1 (8)
Phe	3.7 (4)	5.2 (6)	–	5.2 (6)
Lys	0.9 (1)	1.8 (1)	1.0 (1)	0.8 (1)
His	–	0.7 (1)	–	0.7 (1)
Arg	–	0.3 (1)	–	0.8 (1)
Cys	4.6 (4)	–	–	–

Lys-C-1:

53 60
Asn-Tyr-Gln-Asp-Tyr-Glu-Tyr-Leu-Ile-Asn-Val-Ile-His-

70
-Ala-Phe-Gln-Tyr-Val-Ile-Tyr-Gly-Thr-Ala-Ser-Phe-Phe-Phe-

80 90
-Leu-Tyr-Gly-Ala-Leu-Leu-Leu-Ala-Xaa-Gly-Phe-Tyr-Thr-

100
-Thr-Gly-Ala-Val-Arg-Gln-Ile-Phe-Gly-Asp-Tyr-Lys

Fragment TrpI was cleaved with staphylococcal V8 proteinase. A 10 amino acid-containing peptide (V8-TrpI) had the following sequence:

V8-TrpI:

38 47
Ala-Leu-Thr-Gly-Thr-Glu-Lys-Leu-Ile-Glu

The amino-acid analyses of the fragments T1 (3.6 kDa), T1-TrpI (5.4 kDa), V8-TrpI (1.1 kDa) and Lys-C-1 (5.6 kDa) are shown in Tab. 2. The sequence of TrpI is presented in Fig. 2. Sequences determined by combined gas-liquid chromatography/mass spectrometry[32,33] are underlined.

Fatty acid attachment site in lipophilin

Lipophilin which had been exhaustively reduced and carboxymethylated, was cleaved with papain. Chromatography on Bio-Gel P 100 and high-performance liquid chromatography on Silica gel yielded two peptides, P1 and P2, of about equal size (7.7 kDa) in a ratio of 7:3 as revealed by automated Edman degradation. The two sequences of the peptides as sequenced from the mixture are:

P1:

194 200
Ala-Ser-Ile-Gly-Thr-Leu-Cys-Ala-Asp-Ala-Arg-Met-Tyr-

210
-Gly-Val-Leu-Pro-Trp-Asn-Ala-

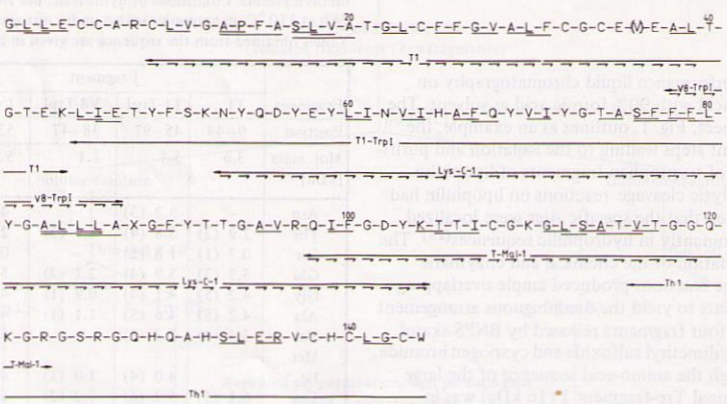

Fig. 2. Amino-acid sequence of HBr/dimethyl sulfoxide fragment I (TrpI), the N-terminal 16-kDa fragment released from reductively carboxymethylated lipophilin with overlapping proteolytic fragments: →, data obtained by automated Edman degradation in the present study; fragments obtained by digestion of total lipophilin with trypsin (T1), thermolysin (Th1) and lysine-specific proteinase (Lys-C-1); T-Mal-1, tryptic fragment of maleoylated lipophilin; V8-TrpI and T1-TrpI, fragments of TrpI-fragment by staphylococcal V8 proteinase and trypsin, respectively. Underlined sequences were determined by gas-liquid chromatography/mass spectrometry.

P2:

^{201}Ala-Asp-Ala-Arg-Met-Tyr-Gly-Val-Leu-Pro-^{210}Trp-Asn-Ala-
-Phe-Pro-Gly-Lys-Val-Cys-^{220}Gly-

A serine residue has been erroneously reported for position 198 in a previous analysis, which corresponded to position 185 in the incomplete sequence with the tentative numbering in this publication[26].

According to Jollès et al.[19] the peptide P1 should contain a fatty acid attachment site. Therefore the phenylthiohydantoin derivatives of Ser^{195} and Thr^{198} were saponified and extracted after acidification, then esterified with methanolic HCl (5%) and the fatty acid esters extracted with hexane. Gas-liquid chromatography of the extract yielded the fatty acid pattern known from the complete peptide fragment of TrpIII[26] (palmitic, stearic and oleic acid in a molar ratio of 1.3 : 1.0 : 1.0). The fatty acid binding site at position 198 could be confirmed by analysing the corresponding phenylthiohydantoin fraction of a thermolysinolytic fragment Th2 with the following sequence:

^{196}Ile-Gly-Thr-Leu-Cys-Ala-Asp-Ala-Arg-Met-Tyr-^{207}Gly-.

A third peptide from maleoylated lipophilin, which was hydrolysed by chymotrypsin, fractionated by Bio-Gel P 60 chromatography and purified by high-performance liquid chromatography on Silica gel Si 100 and Si 60, yielded a 2.2 kDa fragment which was sequenced over 10 cycles:

^{191}Lys-Thr-Ser-Ala-Ser-Ile-Gly-Thr-Leu-^{200}Cys-

Threonine in cycle 8 could hardly be detected but had been identified before in tryptophan fragment III. Saponification and esterification for gas-liquid chromatographic analysis of the hexane extract yielded the fatty acid mixture, identical with that of TrpIII or lipophilin. Phenylthiohydantoin derivatives of Thr^{192}, Ser^{193} and Ser^{195} were free of fatty acids. On this basis the molecular mass of lipophilin is 30 kDa including one covalently bound fatty acid residue. The 276 residues-containing amino acid sequence of lipophilin is given in Fig. 3; only positions 36 and 88 are undetermined. The calculated amino-acid composition as derived from our sequence studies is in good agreement with our experimental data (Tab. 1).

Disulfide bridges

The electrophoretic mobility of lipophilin is strongly dependent on the secondary and tertiary structure and the binding capacity of sodium dodecyl sulfate[25]. Performic acid oxidation abolishes largely the protein conformation. As a consequence, the polyacrylamide gel bands, corresponding to different conformations, collapse in one band with an apparent molecular mass of 30 kDa.

Lipophilin contains 16 cysteine residues. A recent study[34] on the nature of cysteine residues in lipophilin, using the titration with Ellman's reagent and 4-vinylpyridin, supports the estimation of 4.2 to 5.7 half cystines per 100 amino acids, 20% of which are present as cysteines[35–37]. Reductive carboxymethylation of lipophilin in our hands proved to be incomplete and highly dependent on the isolation procedures of lipophilin. During the elucidation of the primary structure of lipophilin so far two out of presumably five disulfide bonds could be located by isolating their linked peptides[25].

a) Trypsin cleavage of maleoylated lipophilin yielded in addition to the above mentioned peptide T-Mal-1 (2.5 kDa) three further fragments T-Mal-2 (17 kDa), T-Mal-3 (10 kDa) and T-Mal-4 (7 kDa). They were purified by Bio-Gel P 150 chromatography and gel permeation high-performance liquid chromatography. Automated Edman degradation indicated that two polypeptides were degraded simultaneously in equal stoichiometry when the 17-kDa fragment was sequenced.

A) ^{1}Gly-Leu-Leu-Glu-Cys-Cys-Ala-Arg-Cys-Leu-^{11}Val-

and

B) ^{205}Met-Tyr-Gly-Val-Leu-Pro-Trp-Asn-Ala-Phe-^{215}Pro-

When this 17-kDa tryptic fragment was oxidized with performic acid, a 10-kDa and 7-kDa peptide arose, the amino-acid sequences of which are identical with those of fragments T-Mal-3 and

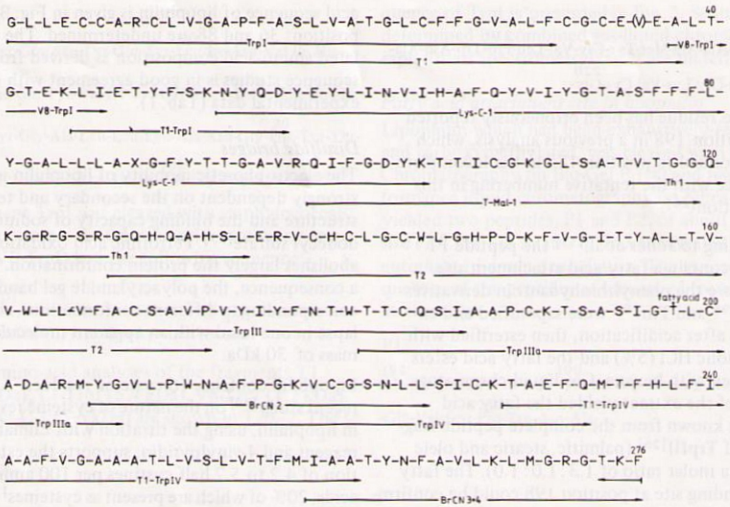

Fig. 3. Complete amino-acid sequence of bovine myelin lipophilin with fatty acid attachment site at Thr^{198}. The horizontal arrows show the sequenced sections of the fragments indicated.
Calculated molecular mass of the apoprotein: 29728 + X Da.

T-Mal-4, respectively. Comparing these two sequences with the complete structure sequence, A corresponds to residues 1 to 11 and B to 205 to 215. Therefore Cys^{219} or Cys^{227} must form a disulfide bond with one of the cysteine residues between residue 5 and 34 at the N-terminus of lipophilin.

b) Enzymatic cleavage of carboxymethylated lipophilin with thermolysin yielded a mixture of peptides from which a 13 kDa large polypeptide was isolated. Again two N-terminal sequences were determined in parallel by Edman degradation, which corresponded to

A) ^{116}Val-Thr-Gly-Gly-Gln-Lys-Gly-Arg-Gly-Ser-Arg-Gly-
-Gln-His-^{129}

and

B) ^{226}Ile-Cys-Lys-Thr-Ala-Glu-Phe-Gln-Met-Thr-Phe-His-
-Leu-Phe-^{239}

Performic acid oxidation cleaved this fragment. This again was proved by high-performance liquid chromatography. Therefore Cys^{227} must form a disulfide bond with a cysteine residue in the center of the polypeptide chain (Fig. 3).

Discussion

The structural analysis of lipophilin (proteolipid protein) from myelin has been made difficult and complicated over the years by the scarce solubility of this, the most hydrophobic protein so far known.

This and our previous reports[25–27] demonstrated that the combination of separation methods allowed the separation and purification of fragments derived from chemical as well as proteolytic cleavage for Edman degradation and gas-liquid chromatography/mass spectrometry.

The following procedures were combined and may prove of general use in the structural analysis of membrane proteins:

- separation of hydrophilic and hydrophobic peptides by precipitation of the latter in acidic aqueous solution
- fractional solubilization of hydrophobic peptides
- gel exclusion chromatography in 90% formic acid
- high-performance liquid chromatography on silica gel of defined pore size of peptides in the range between 20 and 0.5 kDa with 90% formic acid.

Suitable column combinations and rechromatography led to the separation of peptides with differences of 10% in their molecular mass.

Ample overlapping sequences yielded the complete primary structure of lipophilin, which is a strongly hydrophobic 30-kDa protein, Fig. 4. It can be seen in Fig. 4, that the amino-acid sequence of lipophilin is characterized by one short and four long hydrophobic sequences, suitable to integrate into the lipid bilayer. It is obvious from the sequence that the hydrophobic areas are bordered by a charged amino acid or a combination of paired oppositely charged residues. The hydrophobicity of the protein was determined using the hydrophobicity indices of Segrest and Feldman[38]. The position, number of amino acids and hydrophobicity indices of the hydrophobic sequences are summarized in Table 3.

Electron microscopic data suggest a 4.6 nm thickness of the myelin membrane[39]. As compared to the 3 nm bilayer of most of the intracellular membranes[40], 21 amino acids containing hydrophobic sequences are required to span the latter, if an α-helical structure is assumed. About 30 residues long sequences would be required to penetrate the myelin bilayer. Fig. 4 summarizes

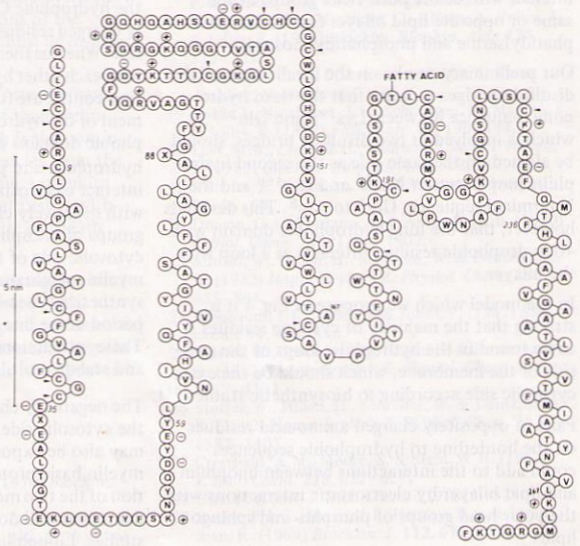

Fig. 4. Suggested integration and orientation of lipophilin in the myelin membrane. The N-terminus is oriented towards the extracytosolic side, the C-terminus towards the cytosolic side of the lipid bilayer of the myelin membrane.

Table 3. Hydrophobicity indices (HI), length and position of hydrophobic amino-acid sequences of lipophilin.

	Position	Number of amino acids	HI
1	9– 34	26	2.31
2	59– 87	29	2.72
3a	151–167	17	3.03
3b	151–190	40	2.36
4	232–267	36	2.35

schematically the proposed integration and orientation of lipophilin in the myelin lipid bilayer. The hydrophobic sequences are linked by hydrophilic, highly charged sequences. The first hydrophilic domain containing 24 amino-acid residues between Glu^{35} and Glu^{58} contains a surplus of negatively charged side chains which could interact with a positive counterpart such as myelin basic protein. The loop containing 72 hydrophilic residues between Xaa^{88} and Lys^{150} is strongly basic and rich in cysteine residues. The basic lysine and arginine side chains could interact with acidic polar head groups of the same or opposite lipid bilayer (sulfatides, phosphatidylserine and phosphatidylinositols).

Our preliminary results on the localization of the disulfide bridges indicate that the third hydrophilic sequence between Lys^{191} and Glu^{231}, which is involved in two disulfide bridges, should be aligned on the same side as the second hydrophilic loop between Xaa^{88} and Lys^{150} and the N-terminal sequence, Gly^{1} to Arg^{8}. This demands however, that the long hydrophobic domain with 40 hydrophobic residues interacts as a loop with the bilayer.

In the model which we propose in Fig. 4 it is striking that the majority of cysteine residues is to be found in the hydrophilic loops of the same side of the membrane, which should be the extracytosolic side according to biosynthetic studies[41].

Pairs of oppositely charged amino-acid residues on the borderline to hydrophobic sequences could add to the interactions between lipophilin and lipid bilayer by electrostatic interactions with the polar head groups of phospho- and sphingolipids.

The suggested model is supported further by our limited proteolysis experiments with non-specific enzymes, e.g. chymotrypsin, thermolysin, papain, subtilisin (not shown here). Their cleavage sites are exclusively in these hydrophilic loops. Also labeling experiments with ^{125}I[42] indicate the partial exposition of lipophilin to the extracytosolic surface of the myelin membrane.

On the basis of experimental as well as on theoretical grounds lipophilin spans the approximately 5 nm lipid bilayer of myelin with long hydrophobic sequences which alternate with strongly charged hydrophilic sequences. It can be regarded as a polytopic membrane protein[43]. Cysteine residues in different hydrophilic areas and loops on the same side of the bilayer linked by disulfide bridges may bring the extra- and intramembranous sequences together to a compact cylinderlike molecule. A further characteristic is the orientation of the positively charged N-terminal sequence and the highly basic hydrophilic loops Xaa^{88} to Lys^{150} and Lys^{191} to Arg^{204}, Lys^{217} to Glu^{231} on one side proposed in our model. The negatively charged hydrophilic sequence between Glu^{43} to Glu^{58} on the other hand, together with the hydrophilic C-terminus carrying three positively charged residues, is positioned on the cytosolic side. Whereas these two sequences may interact with each other by electrostatic interactions and thus contribute further to the condensed arrangement of the hydrophilic and intramembrane hydrophobic domains of lipophilin, the extracytosolic hydrophilic and positively charged areas could interact electrostatically or by hydrogen bonding with negatively charged and other polar head groups of phospho- and sphingolipids of the extracytosolic side of the adjacent bilayer of the myelin membrane wrapped around previously synthesized membranes and thus form the intraperiod dense line seen in electron microscopy[44]. These interactions would strengthen the compact and stabile multilamellar myelin structure.

The negatively charged loop (Glu^{43} to Glu^{58}) on the cytosolic side of the membrane of our model may also be exposed to ionic interactions with myelin basic protein and support the condensation of the two membrane layers to form the main dense line well documented by ultrastructural studies. Lipophilin, therefore, may be regarded as

the main structural element in the myelin membrane linking adjacent bilayers together.

Furthermore, the long-chain fatty acid esterified to Thr^{198}, which is an element of one of the hydrophilic loops of the extracytosolic side must intercalate into the bilayer for thermodynamic reasons. It could, therefore, spread the hydrophilic loop or if the latter is linked by a disulfide bridge also neighbouring loops on the plane of the membrane surface which would enhance ionic interactions.

The empirical rules for the prediction of secondary structures of proteins[45–51] elicit a high degree of ordered structure, about 30% α-helix and 50% β-sheet structure distributed in four long uncharges sequences which are interrupted by random coil areas of different lengths. The latter are those prone for proteolytic cleavage. However, different solvents influence the circular dichroism of lipophilin extremely. Whereas globular proteins interact predominantly with aqueous phases membrane proteins and particularly lipophilin is embedded in the hydrophobic "solvent" of the core of the bilayer. Our circular dichroism measurements showed 80% α-helix structure when chloroform/methanol was the solvent (not shown here). Therefore, predictions of the secondary structure of lipophilin must be regarded with great caution.

The proposed model also could form the basis for the spiral wrapping process of the myelin sheath, the driving force of which is the continuous flow of lipid and protein vesicles from the site of synthesis at the endoplasmic reticulum and Golgi apparatus[41] to sites of fusion with the surface membrane near the site of sheath formation[52].

Whether lipophilin is the molecular equivalent to the ultra structurally documented radial component[53] which has an important structural function by holding the myelin lamellae together at their external surfaces, remains to be proven experimentally.

Literature

1 Boggs, J.M. & Moscarello, M.A. (1978) *Biochim. Biophys. Acta* **515**, 1–21.
2 Eylar, E.H., Brostoff, S., Hashim, G., Caccam, J. & Burnett, P. (1971) *J. Biol. Chem.* **246**, 5770–5784.
3 Folch-Pi, J. & Lees, M. (1951) *J. Biol. Chem.* **191**, 807–817.
4 Folch-Pi, J. & Stoffyn, P.J. (1972) *Ann. N. Y. Acad. Sci.* **195**, 86–107.
5 Lees, M.B., Sakura, J.D., Sapirstein, V.S. & Curatolo, W. (1979) *Biochim. Biophys. Acta* **559**, 209–230.
6 Folch-Pi, J. (1963) in *Brain Lipids and Lipoproteins and the Leucodystrophies* (Folch-Pi, J. & Bauer, H., eds.) p. 18, Elsevier Publ. Co., Amsterdam.
7 Chan, D.S. & Lees, M.B. (1974) *Biochemistry* **13**, 2704–2712.
8 Jollès, J., Nussbaum, J.L., Schöntgen, F., Mandel, P. & Jollès, P. (1977) *FEBS Lett.* **74**, 190–194.
9 Jollès, J., Schöntgen, F., Jollès, P., Vacher, M., Nicot, C. & Alfsen, A. (1979) *Biochem. Biophys. Res. Commun.* **87**, 619–626.
10 Moscarello, M.A., Gagnon, J., Wood, D.D., Antony, J. & Epand, R.M. (1973) *Biochemistry* **12**, 3402–3406.
11 Chan, D.S. & Lees, M.B. (1978) *J. Neurochem.* **30**, 983–990.
12 Sherman, G. & Folch-Pi, J. (1971) *Biochem. Biophys. Res. Commun.* **44**, 157–161.
13 Agrawal, H.C., Randle, Ch.L. & Agrawal, D. (1982) *J. Biol. Chem.* **257**, 4588–4592.
14 Schlesinger, M.J. (1981) *Annu. Rev. Biochem.* **50**, 193–206.
15 Nussbaum, J.L., Rouayrenc, J.F., Mandel, P., Jollès, J. & Jollès, P. (1974) *Biochem. Biophys. Res. Commun.* **57**, 1240–1247.
16 Nussbaum, J.L., Rouayrenc, J.F., Jollès, J., Jollès, P. & Mandel, P. (1974) *FEBS Lett.* **45**, 295–298.
17 Vacher-Leprêtre, M., Nicot, C., Alfsen, A., Jollès, J. & Jollès, P. (1976) *Biochim. Biophys. Acta* **420**, 323–331.
18 Lees, M.B., Chan, D. & Foster, M. (1976) *Trans. Am. Soc. Neurochem.* **7**, 183.
19 Jollès, J., Nussbaum, J.L., Schöntgen, F., Mandel, P. & Jollès, P. (1977) *FEBS Lett.* **74**, 190–194.
20 Lees, M.B., Chao, B.H., Laursen, R.A. & L'Italien, J. (1982) *Biochim. Biophys. Acta* **702**, 117–124.
21 Wachter, E. & Werhan, R. (1979) *Anal. Biochem.* **97**, 56–64.
22 Omenn, G.S., Fontana, A. & Anfinsen, L.B. (1970) *J. Biol. Chem.* **245**, 1895–1902.
23 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 855–864.
24 Nussbaum, J.L., Jollès, J. & Jollès, P. (1982) *Biochimie* **64**, 405–410.
25 Stoffel, W., Schröder, W., Hillen, H. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 1117–1131.
26 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 1397–1407.
27 Crestfield, A.M., Moore, S. & Stein, W.H. (1963) *J. Biol. Chem.* **238**, 622–627.
28 Hirs, C. (1956) *J. Biol. Chem.* **219**, 611–621.
29 Butler, P.J.G., Harris, J.I., Hartley, B.S. & Leberman, R. (1969) *Biochem. J.* **112**, 679–689.

30 Drapeau, G.R. (1977) *Methods Enzymol.* **47**, 189–191.
31 Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91.
32 Nau, H. & Biemann, K. (1976) *Anal. Biochem.* **73**, 154–174.
33 Nau, H. & Biemann, K. (1976) *Anal. Biochem.* **73**, 175–186.
34 Cockle, S.A., Epand, R.M., Stollery, J.G. & Moscarello, M.A. (1980) *J. Biol. Chem.* **255**, 9182–9188.
35 Lees, M.B., Leston, J.A. & Marfey, P. (1969) *J. Neurochem.* **16**, 1025–1032.
36 Nicot, C., Nguyen Le, T., Leprêtre, M. & Alfsen, A. (1973) *Biochim. Biophys. Acta* **322**, 109–123.
37 Nguyen Le, T., Nicot, C., Alfsen, A. & Barratt, M.D. (1976) *Biochim. Biophys. Acta* **427**, 44–56.
38 Segrest, J.P. & Feldman, R.J. (1974) *J. Mol. Biol.* **87**, 853–858.
39 Levine, Y.K. (1975) *Prog. Biophys. Mol. Biol.* **29**, 3–56.
40 Tanford, C. (1978) *Science* **200**, 1012–1018.
41 Colman, D.R., Kreibich, G., Frey, A.B. & Sabatini, D.D. (1982) *J. Cell Biol.* **95**, 598–608.
42 Poduslo, J.F. & Braun, P.E. (1975) *J. Biol. Chem.* **220**, 1099–1105.
43 Blobel, G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1496–1500.
44 Rumsby, M.G. (1978) *Biochem. Soc. Trans.* **6**, 448–462.
45 Chou, P.Y. & Fasman, G.D. (1974) *Biochemistry* **13** 211–222.
46 Fasman, G.D. & Chou, P.Y. (1974) *Biochemistry* **13**, 222–245.
47 Fasman, G.D., Chou, P.Y. & Adler, A.J. (1976) *Biophys. J.* **16**, 1201–1238.
48 Nagano, K. (1973) *J. Mol. Biol.* **75**, 401–420.
49 Nagano, K. (1974) *J. Mol. Biol.* **84**, 337–372.
50 Nagano, K. & Hasegawa, K. (1975) *J. Mol. Biol.* **94**, 257–281.
51 Robson, B. & Suzuki, E. (1976) *J. Mol. Biol.* **107**, 327–356.
52 Rumsby, M.G. (1978) *Biochem. Soc. Trans.* **6**, 448–462.
53 Rumsby, M.G. & Crang, A.J. (1977) *Cell Surf. Rev.* **4**, 247–362.

Prof. Dr. Dr. W. Stoffel, Dr. H. Hillen, Dr. W. Schröder and Dr. R. Deutzmann, Institut für Physiologische Chemie der Universität zu Köln,
Joseph-Stelzmann-Str. 52, D-5000 Köln 41.