Vol. 87, No. 2, 1979 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
March 30, 1979 Pages 619-626

# STRUCTURAL STUDIES OF THE APOPROTEIN OF THE FOLCH-PI BOVINE BRAIN MYELIN PROTEOLIPID : CHARACTERIZATION OF THE CNBr-FRAGMENTS AND OF A LONG C-TERMINAL SEQUENCE

J. Jollès, F. Schoentgen and P. Jollès
Laboratoire des Protéines, Université de Paris V, 45 rue des Saints-Pères, F-75270 Paris Cedex 06

and M. Vacher, C. Nicot and A. Alfsen
Equipe de Recherche des Etats liés moléculaires, C.N.R.S., 45 rue des Saints-Pères, F-75270 Paris Cedex 06, France

Received January 30,1979

SUMMARY. The action of cyanogen bromide on the quite insoluble bovine proteolipid apoprotein allowed the determination of four peptide fragments : two of them constituted a 19 amino acid long C-terminal sequence of the apoprotein. Our results were in favour of the existence of only one subunit presenting a molecular weight closely related to 25,000 for which a schematic representation is given.

To our knowledge, only a small number of sequence studies have so far been devoted to the major brain myelin proteolipid apoprotein. Jollès et al. (1, 2) reported the structural determination of the N-terminal (31 amino acids) and C-terminal (4 amino acids) sequences as well as of 9 tryptic peptides concerning the rat brain apoprotein. The N-terminal sequences of the corresponding human (3) and bovine (4) proteins were also elucidated and it was not possible to detect in the latter any difference with the protein originating from rat brain (1). Moreover the existence of only one type of protein subunit was suggested (4) consequent to the finding of the same N-terminal sequence in the various fractions observed after sodium dodecyl sulfate (SDS) gel electrophoresis.

We decided to extend our sequence studies concerning the Folch-Pi bovine brain myelin apoprotein. The insolubility of the protein and of its peptide fragments in solvents commonly employed during structural stu-

0006-291X/79/060619-08$01.00/0

dies constitutes one of the main difficulties of the present research (2, 5). This paper is dealing with the purification and some structural aspects of the cyanogen bromide (CNBr)-fragments and with the establishment of a long C-terminal sequence.

MATERIALS AND METHODS. Trypsin (EC 3.4.4.4.) was purchased from Worthington. All reagents (analytical grade) were obtained from Merck or Prolabo except those employed for the Sequencer which were purchased from S.D.S. (Marseilles). Sephadex G-25 (coarse) was obtained from Pharmacia and Biogel P 2 (> 400 mesh), P 100 and P 150 (50-100 mesh) from Biorad. The latter were washed with ammonium sulfate and sodium chloride solutions before use in order to avoid electrostatic adsorption of peptides (6).

Apoprotein preparation and free SH group alkylation. The proteolipid was delipidated by light petroleum precipitation and acidic dialyses for three days as described previously (7). At this stage free sulfhydryl group estimation by Ellman's reagent after transfer of the protein in a SDS solution (8) gave a value of 2.4 residues on the 9 half cystine residues found for a molecular weight of 23,500. After one day of neutral dialyses against chloroform/methanol, alkylation has been performed in this medium during 18 h in the conditions of Lees et al. (9). Excess reagent was eliminated by further dialyses at 4°C during 2 days. Transfer of the apoprotein from organic to aqueous solution was performed by the method of Sherman and Folch-Pi (10) in presence of 1 % acetic acid.

Action of cyanogen bromide. The aqueous solution containing the alkylated apoprotein (120 mg) was concentrated by vacuum dialysis, dried and dissolved in 80 % formic acid ; 750 mg CNBr were added and the reaction was allowed to proceed for 24 hours in the dark. Excess CNBr was eliminated under vacuum.

Performic acid oxidation was carried out according to Hirs' procedure (11).

Purification of the CNBr fragments. The CNBr-fragments were separated on Biogel P 100 (150 cm x 3 cm) and Biogel P 150 (200 cm x 3 cm) columns equilibrated with a 90 % formic acid -0.1 M $NH_4OH$ solution. Pure fractions, as determined by the dansyl technique, were pooled, concentrated and precipitated by addition of water to the peptide solution in 90 % formic acid, until a concentration of 20 % formic acid was reached. After centrifugation, the peptide was dissolved in 90 % formic acid ; after dilution with water until a concentration of 70 % formic acid, the peptide was filtered on Sephadex G 25, equilibrated with 50 % acetic acid.

Molecular weight determination. The CNBr-cleavage was performed on the protein delipidated on Biogel P 200 equilibrated in SDS solution (4), reduced and carboxymethylated as previously described (8). The CNBr-fragments were then loaded on a Biogel P 150 column (150 cm x 2 cm) with a 90 % formic acid -0.1 M $NH_4OH$ solution as eluent. Peptides obtained by CNBr-cleavage of cytochrome C were used as molecular weight markers (12).

Structure determination. The amino acid composition of the peptides was determined with a Technicon Autoanalyzer after total hydrolysis.

N-terminal amino-acids were characterized by the dansyl technique or by automated Edman degradation.

Automated Edman degradation was carried out in a Beckman Sequencer Model 890 C by the 1 M quadrol double cleavage method for longer peptides and by the 0.1 M quadrol simple cleavage method in the presence of polybrene (13) for shorter peptides. The thiazolinones were converted into phenylthiohydantoin (PTH)-amino acids by treatment with 20 % trifluoroacetic acid at 80°C during 12 min (14). The PTH-amino acids were characterized by thin-layer chromatography (chloroform-methanol, 90:10, v/v ; pure chloroform), by gas liquid chromatography (Beckman GC 45 chromatograph),with an amino acid Autoanalyzer after regeneration of the free amino acids (150°C, 6 M HCl containing 1/2000 2-mercaptoethanol for 24 h) or by high-performance liquid chromatography (Waters chromatograph, model ALC/GPC-204) according to the procedure of Zeeuws and Strosberg (15).

The C-terminal amino acids of CN-4 were determined by digestion with carboxypeptidases A and B at 37° for different time intervals in 0.1 M ammonium bicarbonate and those of the apoprotein as previously described (4). The digests were then analyzed on a Technicon Autoanalyzer to determine released C-terminal amino acids.

## RESULTS.

Separation of the cyanogen bromide fragments. After CNBr cleavage of the protein three N-terminal amino acids, Gly, Tyr and Ile, were characterized. A partial separation of the CNBr fragments was obtained by gel filtration on Biogel P 100. Two peaks, A and B, were thus characterized.

Two N-terminal amino acids, Gly and Tyr, were found in all the fractions of peak A. Such a behaviour could be explained by the presence of a disulfide bond between two peptides ; this fraction was subjected to performic acid oxidation followed by a filtration on Biogel P 150 (Fig. 1). Peak I contained the non oxidized compound with the two N-terminal amino acids ; peak II corresponded to fragment CN-1 (N-terminal Gly) with an approximate molecular weight of 18,000 and peak III to fragment CN-2 (N-terminal Tyr) with an approximate molecular weight of 4,800.

Peak B contained again two fragments. It was evaporated, suspended in 0.5 M acetic acid and centrifuged ; the precipitate was fragment CN-3 (N-terminal Ile) and the supernatant was characterized as fragment CN-4 (N-terminal Gly) with a slight contamination of CN-3. Its final purification was performed by filtration on Biogel P 2(1.5 cm x 65 cm) equilibrated with 0.5 M acetic acid. Peptide CN-3 was excluded and fragment CN-4 slightly retarded.

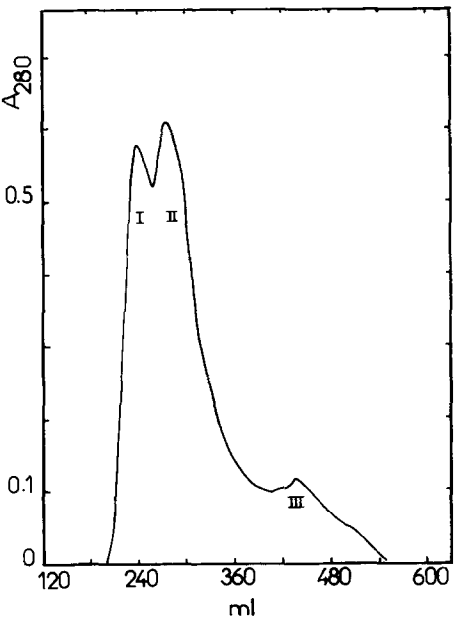



Figure 1 - Filtration on Biogel P 150 (200 cm x 3 cm) of oxidized peak A.

In conclusion four different CNBr-fragments (CN-1 to CN-4) were separated by column chromatography and their homogeneity was ascertained by automated Edman degradation (see below). In order to detect a possible fifth fragment, especially a short sequence, the CNBr-digest of the protein was automatically sequenced prior to any purification. The PTH-amino acids of the first 5 degradation cycles were characterized by high-performance liquid chromatography : they were in accordance with the N-terminal sequences of the sole fragments CN-1 to CN-4 (see below). Furthermore the eluate of the Biogel P 100 column corresponding to the area of low molecular weight material(s) was hydrolyzed and analyzed : no short peptide could be characterized.

Sequence determination of the CNBr-fragments.

Fragment CN-1. Automated Edman degradation permitted the establishment of the N-terminal sequence up to the 23rd amino acid residue : it was identical to the N-terminal sequence of the undigested apoprotein (4).

Thus CN-1 constituted the N-terminal part of the proteolipid apoprotein (Table 1). Residue N° 5 was again characterized as cysteine (4).

Fragment CN-2. The 18 first amino acids were automatically determined (Table 1). In order to separate CN-2 from CN-1 after CNBr-cleavage, it was necessary to carry out a performic acid oxidation. This result indicated the existence of at least one disulfide bond between the two fragments.

Fragment CN-3. Fragment CN-3 is a tridecapeptide : $Asp_1$, $Thr_1$, $Ala_3$, $Met_1$, $Val_1$, $Ile_1$, $Leu_2$, $Tyr_1$, $Phe_1$, $Lys_1$. The first 10 amino acids were automatically sequenced (Table 1). After tryptic digestion, the dipeptide Leu-Met was characterized by usual techniques. The structure of CN-3 is indicated in Table 2.

Fragment CN-4. Fragment CN-4 is a hexapeptide : $Thr_1$, $Gly_2$, $Phe_1$, $Lys_1$, $Arg_1$. Its structure was automatically established (Table 1). Furthermore digestion experiments with carboxypeptidases A and B corroborated the establishment of the C-terminal sequence Lys-Phe. After digestion with carboxypeptidase A, Phe has already been characterized as the C-terminal amino acid of the apoprotein (4). When the latter was submitted to the action of carboxypeptidases A and B, only amino acids consistent with the sequence of fragment CN-4 were detected.

CONCLUSION. Our results concerning the CNBr-fragments of bovine proteolipid apoprotein are summarized in Table 2. In a previous study devoted to the rat apoprotein, Jollès *et al.* (2) characterized the tryptic tetrapeptide : Leu-Met-Gly-Arg which was also found by Chan *et al.* (5) in the bovine proteolipid apoprotein. Thus we suggest that the above mentioned tryptic tetrapeptide allows the alignment CN-3 → CN-4 : it constitutes a 19 amino acid residues long C-terminal sequence of the proteolipid apoprotein preceded by a methionine residue. It is worth mentioning that among

TABLE 1

Automatic Edman degradation of CNBr-fragments from bovine proteolipid apoprotein.

| Cycle | CN-1 | | | CN-2 | | | CN-3 | | | CN-4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | a | b | c | a | b | c | a | b | c |
| 1 | Gly | Gly(19) | Gly(19) | Tyr | Tyr(25) | Tyr(12) | Leus | Leus(33) | Ile(29) | Gly | Gly(20) | Gly(18) |
| 2 | Leus | Leus(32) | Leu(29) | Gly | Gly(10) | Gly(20) | Ala | Ala(24) | Ala(25) | Arg | | Arg(6) |
| 3 | Leus | Leus(33) | Leu(32) | Val/Phe | Val(25) | Val(22) | Ala | Ala(27) | Ala(23) | Gly | Gly(15) | Gly(16) |
| 4 | Glu | | Glu(27) | Leus | Leus(20) | Leu(20) | Thr | Thr(+) | | Thr | Thr(+) | |
| 5 | Cys° | | SCM(3) | Pro | Pro(10) | Pro(8) | Tyr | Tyr(8) | Tyr(6) | Lys | | Lys(14) |
| 6 | * | | | Lys/Trp | Trp(7) | | Asn | | Asp(3) | Phe | Phe(3) | |
| 7 | Ala | Ala(29) | Ala(19) | Asn | | Asp(13) | Val/Phe | Phe(5) | Phe(4) | | | |
| 8 | | | Arg(6) | Ala | Ala(14) | Ala(16) | Ala | Ala(3) | Ala(2) | | | |
| 9 | * | | | Val/Phe | Phe(9) | Phe(15) | Val/Phe | Val(3) | Val(3) | | | |
| 10 | Leus | Leus(14) | Leu(15) | Pro | Pro(6) | Pro(5) | Leus | Leus(2) | Leu(1) | | | |
| 11 | Val/Phe | Val(14) | Val(12) | Gly | Gly(6) | Gly(8) | | | | | | |
| 12 | Gly | Gly(8) | Gly(9) | Lys/Trp | | Lys(7) | | | | | | |
| 13 | Ala | Ala(8) | Ala(9) | Val/Phe | Val(7) | Val(7) | | | | | | |
| 14 | Pro | Pro(6) | Pro(4) | | | | | | | | | |
| 15 | Val/Phe | Phe(4) | Phe(7) | Gly | Gly(2) | Gly(3) | | | | | | |
| 16 | Ala | Ala(7) | Ala(8) | Ser | Ser(4) | | | | | | | |
| 17 | Ser | Ser(8) | | Asn | | Asp(3) | | | | | | |
| 18 | Leus | Leus(7) | Leu(5) | Leus | Leus(3) | Leu(4) | | | | | | |
| 19 | Val/Phe | Val(6) | Val(7) | | | | | | | | | |
| 20 | Ala | Ala(6) | Ala(5) | | | | | | | | | |
| 21 | Thr | Thr(+) | | | | | | | | | | |
| 22 | Gly | Gly(4) | Gly(2) | | | | | | | | | |
| 23 | Leus | Leus(4) | Leu(2) | | | | | | | | | |

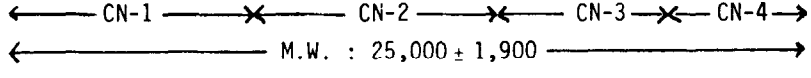
a, thin layer chromatography
b, gas-liquid chromatography, results give the percentage
c, amino acid determined with an Autoanalyzer after regeneration, results give the percentage
°, Cys, characterized as phenylthiohydantoin-S-carboxamidomethylcysteine (alkylation without reduction)
*, see Ref. 4 : Cys was determined only when the apoprotein was reduced before alkylation.
Leus, Leu or Ile, characterization possible by amino acid analysis ; SCM, S-carboxymethylcysteine.

TABLE 2

Structural data concerning the CNBr-fragments isolated from bovine brain myelin proteolipid apoprotein. Schematic representation of the apoprotein.

| CNBr-fragments | M.W.* | Structure | Observation |
|---|---|---|---|
| CN-1 | 18,000 ± 1,500 | Gly-Leu-Leu-Glu-Cys-Cys-Ala-Arg-Cys-Leu-Val-Gly-Ala-Pro-Phe-Ala-Ser-Leu-Val-Ala-Thr-Gly-Leu-....-Met | N-terminal sequence of the apoprotein |
| CN-2 | 4,800 ± 400 | Tyr-Gly-Val-Leu-Pro-Trp-Asn-Ala-Phe-Pro-Gly-Lys-Val- X -Gly-Ser-Asn-Leu-.....-Met | |
| CN-3 | 1,500 (calculated 1,453) | Ile-Ala-Ala-Thr-Tyr-Asn-Phe-Ala-Val-Leu-Lys-Leu-Met | Alignment CN-3 → CN-4 established (see text) |
| CN-4 | 700 (calculated 664) | Gly-Arg-Gly-Thr-Lys-Phe | C-terminal sequence of the apoprotein |

Schematic representation of the apoprotein (alignment of the CNBr-fragments) :




* M.W. determined by filtration on Biogel (see Methods).

the last nine residues of this fragment two lysine and one arginine residues occur ; thus the C-terminal sequence is highly positively charged as previously observed for the whole protein (7).

The determination of only one long C-terminal sequence as well as of only one N-terminal sequence (2, 4) does not seem to be in favour of the suggestion concerning the existence of two non identical subunits (5). Thus according to the number of the CNBr-fragments and to their size, the apoprotein subunit should possess a molecular weight not lower than 23,000 as suggested previously (2) : its schematic representation is indicated in Table 2.

ACKNOWLEDGEMENTS. This research was supported in part by the C.N.R.S. (E.R. 102 and E.R. 64) and the I.N.S.E.R.M. (group U-116, FRA N° 27 and CL 78.4.0413). The excellent technical assistance (automated Edman degradation) of Mr Ly Quan Le is gratefully acknowledged.

REFERENCES

1. Nussbaum J.L., Rouayrenc J.F., Mandel P., Jollès J. and Jollès P. (1974) Biochem. Biophys. Res. Commun. 57, 1240-1247.
2. Jollès J., Nussbaum J.L., Schoentgen F., Mandel P. and Jollès P. (1977) FEBS Lett. 74, 190-194.
3. Nussbaum J.L., Rouayrenc J.F., Jollès J., Jollès P. and Mandel P. (1974) FEBS Lett. 45, 295-298.
4. Vacher-Leprêtre M., Nicot C., Alfsen A., Jollès J. and Jollès P. (1976) Biochim. Biophys. Acta 420, 321-331.
5. Chan D.S. and Lees M.B. (1978) J. Neurochem. 30, 983-990.
6. Mehl E. and Halaris A. (1970) J. Neurochem. 17, 659-668.
7. Nguyen Le.T., Nicot C., Alfsen A. and Barrat M.D. (1976) Biochim. Biophys. Acta 427, 44-56.
8. Nicot C., Nguyen Le.T., Leprêtre M. and Alfsen, A. (1973) Biochim. Biophys. Acta 322, 109-123.
9. Lees M.B., Leston J.A. and Marfey P. (1969) J. Neurochem. 16, 1025-1032.
10. Sherman G. and Folch-Pi J. (1970) J. Neurochem. 17, 597-605.
11. Hirs C.H.W. (1956) J. Biol. Chem. 219, 611-621.
12. Swank R.T. and Munkres K.D. (1971) Anal. Biochem. 39, 462-477.
13. Klapper D.G., Wilde III C.E. and Capra J.D. (1978) Anal. Biochem. 85, 126-131.
14. Sladic-Simic D., Kleinschmidt T. and Braunitzer G. (1977) Z. phys. Chem. 358, 591-594.
15. Zeeuws R. and Strosberg A.D. (1978) FEBS Lett. 85, 68-72.