Proteolysis of the neurofilament 68 kDa protein explains several previously described brain proteins of unique composition and high acidity

Klaus Weber and Norbert Geisler

Max Planck Institut für Biophysikalische Chemie, 3400 Göttingen, FRG

Received 13 October 1983

Neurofilaments follow the structural principles of non-neuronal intermediate filaments but contain additional sequences which are carboxyterminally located and increase in length between triplet proteins (68 kDa, 160 kDa and 200 kDa). The tailpiece domain has been sequenced in the case of the porcine 68 kDa protein. It has a unique amino acid composition. Within 106 residues there are only 12 different amino acid types, and glutamic acid accounts for 46% of the sequence. Examination of the literature on highly acidic brain proteins leads us to the proposal that microglutamic acid-rich protein, Glu-50, macroglutamic protein, as well as some unusual components of the S100 class, are most likely proteolytic degradation products of the neurofilament 68 kDa protein.

Acidic brain protein *Glutamic acid-rich Neurofilament* *Proteolysis S100 Protein* *Intermediate filament*

1. INTRODUCTION

Extensive sequence work has established that all 5 subclasses of intermediate filamets, as well as the wool α-keratins, reveal a common structural principle. Some 310 residues form a highly α-helical domain able to build interpolypeptide double-stranded coiled-coils in agreement with the α-type diffraction pattern (review [1]). Within this rod domain sequence identity values for the different proteins vary between 30 and 70%. The rod is flanked by two non-α-helical domains which vary greatly among the different filament proteins both in sequence and in length. Neurofilaments are restricted to neurones. The three mammalian triplet proteins (68, 160 and 200 kDa) carry extra sequences which are located exclusively at the carboxyl end. These tailpieces show unique amino acid compositions and increase in length among the three proteins [2]. We have recently sequenced the tailpiece of the 68 kDa protein and defined two structurally distinct domains. The carboxyterminal array of 106 residues is very acidic due to the presence of 47 residues of glutamic acid [2]. We here show that four very acidic brain proteins described in detail by others [3–6] can be explained as proteolytic derivatives of this unique region of the 68 kDa neurofilament protein.

2. RESULTS AND DISCUSSION

We first consider microglutamic acid-rich protein (I) and Glu-50 (II) isolated from bovine and human brain, respectively [3,4]. For simplicity we abbreviate these proteins as I and II. Both proteins have been described as having a polypeptide molecular mass of some 10 kDa and have been well characterized both by amino acid composition and by the determination of their aminoterminal sequence [3,4]. As shown in fig.1 the first 12 established residues of bovine I [4] can be perfectly aligned with residues −82 to −71 of the porcine neurofilament 68 kDa tailpiece. The first 6 known residues of II [3] fit to the extent that only a serine

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

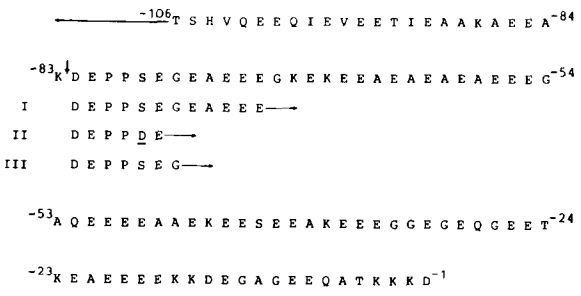

Fig.1. Comparison of the aminoterminal sequences [3–5] reported for three uniquely glutamic acid-rich brain proteins (I–III) with the established sequence of the carboxyterminal 106 residues of the major neurofilament triplet protein 68 kDa [2]. For convenience of comparison the carboxyterminal residue is given position −1. Note the alignment of the aminoterminal residues of the three proteins with the residues following lysine (position −83). In the case of the human protein II there may be a species-specific amino acid exchange (underlined). The vertical arrow indicates the proposed site of cleavage.

to an aspartic acid exchange has to be permitted as a species-specific event upon transition from the bovine or porcine protein to its human counterpart. Corresponding recalculation of the published amino acid compositions [3,4] (table 1) shows a very convincing fit if we assume that I and II cover the carboxyterminal 82 residues [2] of the neurofilament 68 kDa tailpiece array. The resulting fragments are unique in amino acid composition since only 8 residue types (glutamic acid, alanine, lysine, glycine, aspartic acid, threonine, serine and proline) account for the molecules. Since the content of 51% glutamic acid is accounted for the molecules. Since the content of 51% glutamic acid is accounted for by 39 glutamic acid and 3 glutamine residues [2] a uniquely glutamic acid-rich domain is seen. Because this contains by sequence studies only 10 lysines but an additional 3 aspartic acids [2] a very low isoelectric point can be predicted, and direct measurements gave values of about 3.9 [3,4]. In addition carboxypeptidase A plus B digestion of II indicated some two to three lysines [3] at the carboxyterminal end, which is entirely compatible with the neurofilament sequence (fig.1) if the ultimate aspartic acid were also removed by previous proteolytic processing. The reported oligomeric character of II [3] fits the established dimeric state of the carboxyterminal 106 residues of the 68 kDa tail domain [2] and the high resistance to thermolysin and moderate resistance to trypsin noted before were also experienced in our studies on the 68 kDa tailpiece. Equally the sensitivity to V8-protease digestion holds for I, II and the neurofilament tailpiece domain. Thus all data on I and II are consistent with the hypothesis that these proteins could arise by a defined proteolytic break past lysine residue −83 in the carboxyterminal domain of the major neurofilament protein [2].

Table 1

Amino acid composition of three very acidic brain proteins (I–III) recalculated for the 82 residues corresponding to the carboxyterminal array of the neurofilament 68 kDa protein (IV)

	I	II	III	IV
Glu	41	41	41	42
Ala	14	10	11	12
Lys	10	9	9	10
Gly	7	9	9	9
Asp	4	4	3	3
Thr	2	2	2	2
Ser	2	2	2	2
Pro	2	2	2	2

Original compositions have been given before: I [4], II [3], III [5]. The composition of IV is taken from the established sequence [2] see also fig.1. All other amino acids are present only in trace amounts [3,4]

For the third related protein [5], macroglutamic protein from porcine brain (III), there is again a perfect fit for the reported 7 aminoterminal residues (fig.1) and for a recalculated amino acid composition assuming the same derivation as discussed above. We note, however, that III was assumed to have a polypeptide molecular mass of 57 kDa rather than 10 kDa [5]. We question this value because of two reasons. First, in gel filtration on Sephadex G-75 III eluted later than the calmodulin present in an earlier purification step [5]. Since the latter protein has a molecular mass of 16 kDa a direct discrepancy in molecular mass assignment is revealed. Second, the sample used in phosphate–SDS gel electrophoresis was heated for 10 min at 60°C in 8 M urea [5]. This treatment is expected to yield a heavily carbamylated protein, which because of its nearly exclusive negative charges cannot be expected to show a normal mobility in gel electro-

phoresis. Alternatively we cannot exclude that III may indeed have a higher molecular mass and arises not from 68 kDa protein but rather from one of the higher molecular mass neurofilament triplet proteins, assuming that they contain analogous sequences [2].

Finally we notice that the very acidic components F_3C and F_4C in a collection of presumptive S100 brain proteins [6] do not show the typically high content of phenylalanine characteristic of S100 proteins established by sequence [7]. The amino composition of these two components would be again consistent with their derivation from neurofilament 68 kDa protein. In conclusion all the proteins described above, peculiar because of their high acidity and amino acid composition, could readily arise by a restricted proteolytic trimming of a major neuronal protein, the neurofilament 68 kDa protein. Although we cannot decide if this cleavage occurs artificially during tissue fractionation or already in situ we note that proteolytic processing of neurofilaments has been postulated to be a normal intracellular event [8,9].

REFERENCES

[1] Geisler, N. and Weber, K. (1982) EMBO J. 1, 1649–1656.
[2] Geisler, N., Kaufmann, E., Fischer, S., Plessmann, U. and Weber, K. (1983) EMBO J. 2, 1295–1302.
[3] Nomata, Y., Watanabe, T. and Wada, H. (1983) J. Biochem. 93, 825–831.
[4] Isobe, T., Ishioka, N., Kadoya, T. and Okuyama, T. (1982) Biochem. Biophys. Res. Commun. 105, 997–1004.
[5] Ishioka, N., Isobe, T., Okuyama, T., Numata, Y. and Wada, N. (1980) Biochim. Biophys. Acta 625, 281–290.
[6] Starostina, M.V., Malup, T.K. and Sviridov, S.M. (1981) J. Neurochem. 36, 1904–1915.
[7] Isobe, T. and Okuyama, T. (1978) Eur. J. Biochem. 89, 379–388.
[8] Calvert, R. and Anderton, B.H. (1982) FEBS Lett. 145, 171–175.
[9] Calvert, R. (1982) Nature 300, 580–581.