Brain micro glutamic acid-rich protein is the C-terminal endpiece of the neurofilament 68-kDa protein as determined by the primary sequence

Toshiaki Isobe and Tsuneo Okuyama

Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, Setagaya, Tokyo, Japan 158

Received 24 January 1985

The amino acid sequence of bovine brain micro glutamic acid-rich protein was determined by analysis of tryptic and *Trimeresurus flavoviridis* protease peptides of the molecule. The protein comprised 82 amino acid residues and has an M_r of 8992. The established sequence was highly homologous (90% identity) to the sequence of C-terminal 82 residues of the neurofilament 68-kDa protein from porcine spinal cord; there are differences of 8 residues which could be species-specific amino acid substitutions. This indicates that the micro glutamic acid-rich protein may arise by a restricted proteolysis of the neurofilament 68-kDa protein, with the break occuring toward the C-terminus.

Glutamic acid-rich protein *Amino acid sequence* *Neurofilament* *Proteolysis* *Calcium activated protease* *Brain*

1. INTRODUCTION

Micro glutamic acid-rich protein (MGP) is an acidic (pI = 3.9) low molecular mass (10-kDa) protein found in bovine brain tissue [1]. This protein has a characteristic amino acid composition. There are only 8–10 amino acid types, with no 'hydrophobic' amino acids, and glutamic acid/glutamine accounts for 50% of the total composition. Nomata et al. [2] isolated the human equivalent of this protein and demonstrated immunochemically that it is distributed in brain tissue. However, no biological activities have been found for this protein so far.

Recently Geisler et al. [3] have presented the partial amino acid sequence of one of the three neurofilament triplet proteins, often referred to as the 68-kDa protein. They defined two structurally distinct domains. The N-terminal domain shares a number of structural characteristics in common with non-neuronal intermediate filament proteins, while the C-terminal domain has a structure which is rather specific to neurofilament protein. It has been postulated that the N-terminal domain allows the 68-kDa protein to polymerize into the filament structure and the C-terminal domain provides a scaffold for interaction with other cellular components [3]. Weber and Geisler [4] also noticed that the structure of the C-terminal domain resembled MGP in terms of amino acid composition and N-terminal sequence.

We report here the amino acid sequence of MGP isolated from bovine brain cytosolic fraction. The sequence suggests that MGP arise by a restricted proteolysis of the neurofilament 68-kDa protein, where the cleavage takes place at a Lys-Asp bond in the C-terminal domain.

Abbreviations: MGP, micro glutamic acid-rich protein; HPLC, high performance liquid chromatography

2. MATERIALS AND METHODS

MGP was purified from bovine brain extract as described in [1]. The final purification was achieved by preparative HPLC performed on a column (18 × 200 mm) of Hitachi gel 3013 resin (Hitachi, Tokyo) using an acetonitrile-trifluoro-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

acetic acid solvent system [5]. Digestions of MGP (0.5–1 mg) with L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone-treated trypsin (Millipore, NJ) or with a neutral protease from *Trimeresurus flavoviridis* were carried out in 50 mM ammonium bicarbonate (pH 8.1) for 0.5–5 h at an enzyme-substrate ratio of 1:50 (w/w). The digests were subsequently applied to a column (4.6 × 250 mm) of Finepak-sil ODS (Japan Spectroscopic, Tokyo) or of Develosil ODS (Nomura Chemicals, Aichi) and peptides eluted by a mobile phase gradient of acetonitrile in 0.1% (v/v) trifluoroacetic acid. Sequence analyses including amino acid analysis, carboxypeptidase digestion, hydrazinolysis, and Edman degradation followed by HPLC identification of the resulting phenylthiohydantoin amino acids were performed as in [6,7] with modifications to adapt for micro-scale experiments [8,9].

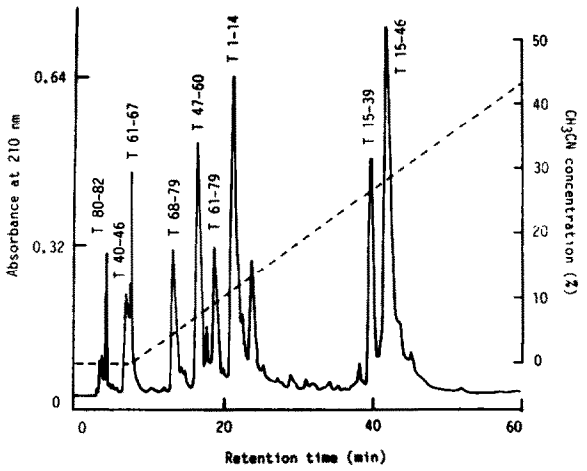

Fig.2. Separation of the tryptic peptides of MGP by reversed phase HPLC. The tryptic peptides obtained after 5 h digestion of MGP (100 nmol) were applied to a column (4.6 × 250 mm) of Develosil ODS (3 μm) and eluted as described in section 2.

3. RESULTS AND DISCUSSION

The amino acid sequence of MGP was deduced from tryptic and *Trimeresurus* protease peptides of the molecule. Fig.1 shows the sequence of MGP and the peptides used to establish the sequence.

The tryptic peptides were obtained from two batches of digests which differed in the extent of cleavage. By a 30-min digestion, the tryptic cleavage of the MGP molecule occurred preferentially at Lys-60 and Lys-79 and generated 3 fragments, T 1–60, T 61–79, and T 80–82. Further extended digestion (5 h) produced an additional 7 peptides, T 1–14, T 15–39, etc. (fig.2), while T 80–82 and a portion (30%) of T 61–79 remained uncleaved. Sequential Edman degradation of these tryptic peptides followed by quantitative examination of the resulting phenylthiohydantoin amino acids by HPLC, and carboxypeptidase digestions supplementarily used for the C-terminal analysis enabled determination of most of the MGP sequence shown in fig.1. The sequence of residues 33–37 was hardly evaluated from the

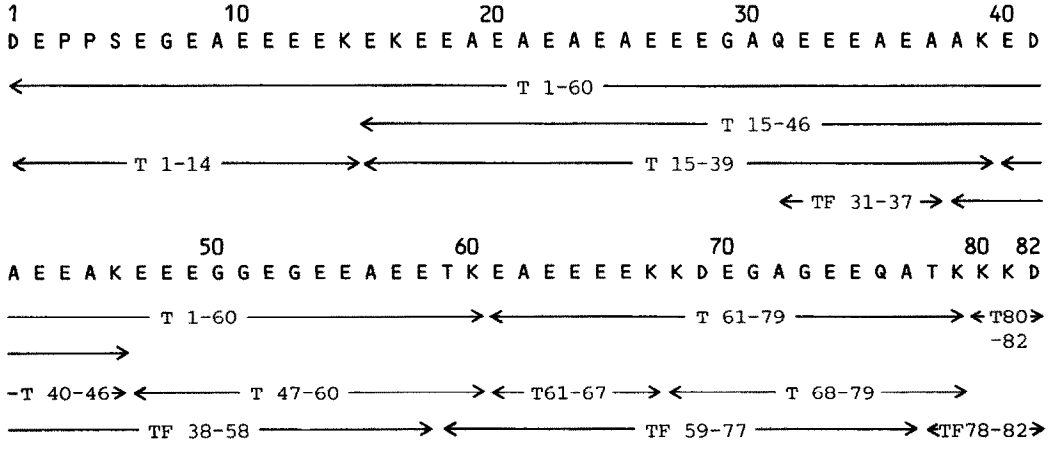

Fig.1. The amino acid sequence of bovine brain MGP and peptides needed for the sequence determination. T refers to tryptic peptides and TF to *T. flavoviridis* protease peptides.

analyses of T 15–39 and T 15–46 due to a low yield of corresponding phenylthiohydantoin amino acids. Therefore it was determined by the analysis of a *Trimeresurus* protease peptide TF 31–37. The total sequence of MGP was established from the sequence of these peptides and the sequence of 3 overlapping peptides TF 38–58, TF 59–77, and TF 78–82 (fig.1), which was derived from the *Trimeresurus* protease digest of MGP. The established sequence (fig.1) accounted for the reported amino acid composition recalculated from [1]; Asp 3.6 (4), Thr 1.7 (2), Ser 0.9 (1), Glu + Gln 42.0 (41), Pro 2.5 (2), Gly 7.0 (7), Ala 15.8 (15), and Lys 9.5 (10), where numbers in parentheses are the numbers of residues computed from the sequence. Based on the sequence, MGP is composed of 82 residues and has an M_r of 8992. This value is in close agreement with that estimated by gel electrophoresis analysis [1].

Earlier, we presented the sequence of N-terminal 18 residues [1]. This sequence, based on Edman degradation of an intact MGP, differs from the final sequence by the assignment of residues 14 and 16; we have previously located glutamic acid at both these positions instead of lysines in fig.1. Here, using proteolytic fragments, these residues have easily been identified because residue 14 locates at the C-terminus of T 1–14 (it was a site of tryptic cleavage) and residue 16 locates next to the N-terminus of T 15–39 and T 15–46. The results were clear and conclusive. Reinspection of the previous Edman degradation data has also revealed that there were increases in lysine residues at steps 14 and 16. We have therefore concluded that lysines are at these positions. It is likely that the highly repetitive sequence of glutamic acid at positions 9–13 caused an unexpected degree of carry-over toward the subsequent Edman degradation steps [1] and led us to an erroneous interpretation of the data for these steps.

Weber and Geisler [4] noticed that the amino acid composition and N-terminal sequence of MGP resembled that of the C-terminal region of the neurofilament 68-kDa protein, and suggested the relationship between these proteins. A comparison of the established MGP sequence and the equivalent sequence of the 68-kDa protein demonstrates, in fact, a striking degree of homology between the sequence; there are differences of only 8, among 82, amino acid residues (fig.3). Considering that these differences could be amino acid substitutions due to divergence of the two species, it is likely that MGP arises by a restricted proteolytic cleavage of the neurofilament 68-kDa protein as suggested by Weber and Geisler [4]. We note that all the observed differences are compatible to single nucleotide-base mutations in their triplet codons.

The proteolytic processing of neurofilaments has been observed in previous reports [10] and is now considered to be a normal intracellular event somehow related to cytoplasmic transport in neurons [11,12]. It occurs in the presence of calcium [10] and there is some evidence that calcium-activated protease is responsible for this process [13,14]. However, details of the cleavage

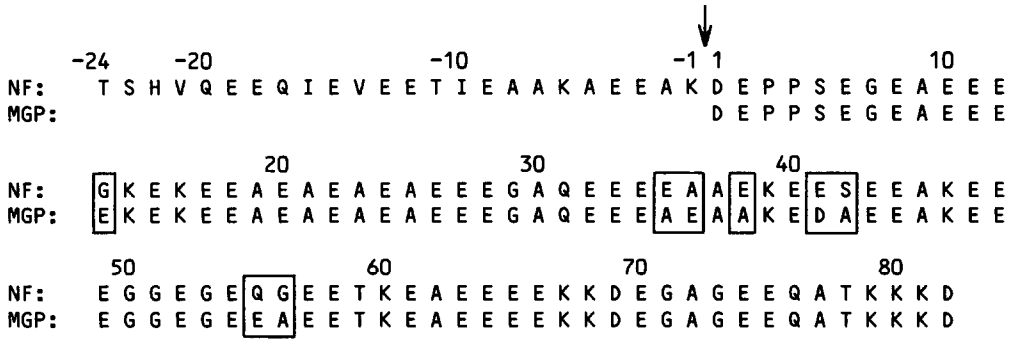

Fig.3. Comparison of the established sequence of bovine brain MGP with the sequence of the C-terminal 106 residues of the neurofilament 68-kDa protein (NF) from porcine spinal cord [3]. The difference lies in 8 positions (boxed). The vertical arrow indicates the cleavage site. The amino acid residues are numbered beginning with the N-terminus of MGP and the preceding residues are designated by negative numbers.

mechanism and specificity, or of the resulting morphological changes in neurofilament are unknown. In order to generate the MGP molecule, the cleavage of the 68-kDa protein should have occurred at a Lys(−1)-Asp(1) bond (fig.3). The sequence of MGP suggests that this specificity is considerably high because 10 lysine residues, including the cleavage sequence Lys-Asp, still occur in the MGP sequence (fig.1). Although it has so far remained unclear whether this cleavage takes place in vivo or during the MGP preparation we note that the procedure for purifying MGP includes EDTA, an inhibitor of putative calcium activated protease, at a relatively high concentration (1 mM). In this procedure we have constantly obtained MGP from the brain cytosolic fraction at a level of 1 mg/kg wet tissue [1].

Based on the structural comparison among intermediate filament proteins, Geisler et al. [3] postulated that the C-terminal domain of the 68-kDa protein, including the MGP sequence, determines 'neurofilament specificity' by providing a highly charged scaffold suitable for interaction with other cellular components. From this viewpoint, the proteolytic removal of MGP may be a regulatory mechanism of neurofilament architecture as it should result in an altered interaction and cause morphological changes in neurofilaments. Alternatively, it is possible that the free MGP molecule plays a unique biological role in vivo.

ACKNOWLEDGEMENTS

The authors wish to thank Drs T. Sato and Y. Yamakawa of the National Institute of Health, Tokyo, for donation of the neutral protease from *T. flavoviridis*. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

REFERENCES

[1] Isobe, T., Ishioka, N., Kadoya, T. and Okuyama, T. (1982) Biochem. Biophys. Res. Commun. 105, 997–1004.
[2] Nomata, Y., Watanabe, T. and Wada, H. (1983) J. Biochem. 94, 1269–1278.
[3] Geisler, N., Kaufmann, E., Fischer, S., Plessmann, U. and Weber, K. (1983) EMBO J. 2, 1295–1302.
[4] Weber, K. and Geisler, N. (1983) FEBS Lett. 164, 129–131.
[5] Isobe, T., Kurosu, Y., Fang, Y.-I., Ishioka, N., Takai, N. and Okuyama, T. (1984) J. Liq. Chromatog. 7, 1101–1115.
[6] Isobe, T. and Okuyama, T. (1978) Eur. J. Biochem. 89, 379–388.
[7] Isobe, T. and Okuyama, T. (1981) Eur. J. Biochem. 116, 79–86.
[8] Miyachi, H., Fan, Y.-I., Kametani, F., Isobe, T., Shinoda, T. and Okuyama, T. (1983) Seikagaku 55, 803.
[9] Ichimura, T., Miyachi, H., Isobe, T. and Okuyama, T. (1984) Seikagaku 56, 812.
[10] Norton, T. and Goldman, J.E. (1980) in: Proteins of the Nervous System (Bradshaw, R.A. and Schneider, D.M. eds) 2nd edn, pp.301–329, Raven, New York.
[11] Calvert, R. and Anderton, B.H. (1982) FEBS Lett. 145, 171–175.
[12] Calvert, R. (1982) Nature 300, 580–581.
[13] Schlaepfer, W.W. and Hasler, M.B. (1979) Brain Res. 168, 299–309.
[14] Nixon, R.A., Brown, B.A. and Marotta, C.A. (1981) J. Cell Biol. 94, 150–158.