Volume 210, number 2, 157–160 FEB 04331 January 1987

Molecular cloning of cDNA coding for human PGP 9.5 protein

A novel cytoplasmic marker for neurones and neuroendocrine cells

Ian N.M. Day and R.J. Thompson

Department of Clinical Biochemistry, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QR, England

Received 22 October 1986

The co-ordinate sequencing of the human neuronal and neuroendocrine marker protein PGP 9.5 and its cDNA is described. The cDNA encodes the complete protein (212 amino acids), and the 340 nucleotide 3′-noncoding region including the polyadenylation signal, indicating an mRNA slightly larger than 1 kb in size. Protein sequencing of 50% of PGP 9.5 confirms the deduced protein sequence.

Protein gene product 9.5; Neuron; Neuroendocrine system; cDNA sequence; (Brain)

1. INTRODUCTION

Protein gene product (PGP) 9.5 is a novel neurone-specific protein originally discovered by high-resolution two-dimensional mapping of soluble proteins from different human organs [1,2]. Immunohistochemistry has shown that PGP 9.5 is widely distributed in central and peripheral neurones [3], in the retina [4], and in cells of the diffuse neuroendocrine system (DNES) [4]. This distribution closely parallels that of neurone-specific enolase (NSE), which is the only other general cytoplasmic marker known for these cell types [5]. The structure, properties, and evolutionary distribution of PGP 9.5 have recently been reviewed [6].

To date we have been concerned with the detection of human PGP 9.5 as a disease marker using polyclonal (e.g. [7]) and monoclonal [6,8] antibodies: here we describe the co-ordinate characterization of human PGP 9.5 and its cDNA to provide an additional probe in the analysis of the developmental and functional biology and pathology of neurones and neuroendocrine cells.

Correspondence address: I.N.M. Day, Department of Clinical Biochemistry, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QR, England

2. MATERIALS AND METHODS

Trypsin and chymotrypsin were from Worthington, acetonitrile from Rathburn Chemicals, dabsylisothiocyanate from Fluka, [γ-^{32}P]ATP and [α-^{32}P]dCTP from Amersham, nitrocellulose from Schleicher and Schuell. Other reagents and enzymes were from Sigma, Boehringer, Mannheim and New England Biolabs.

2.1. *Analysis of pure human PGP 9.5 protein*

N-terminal analysis, and a variety of digestions analysed by manual and automated gas-phase sequential analysis were performed. Briefly the most useful cleavages and resolutions of digests obtained were: (i) partial cyanogen bromide cleavage of unreduced PGP 9.5, then direct sequencing; (ii) tryptic digestion (1:10, w/w; 3 h at 37°C; 0.1%

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

ammonium bicarbonate) of performic acid oxidized PGP 9.5 followed by 3 MM paper electrophoresis (pyridine/acetic acid/water, 100:4:900, v/v) and then perpendicular chromatography (isoamyl alcohol/pyridine/water, 7:7:6, v/v) with detection by partial fluorescamine staining; (iii) chymotryptic digestion (conditions as (ii)) of fully carboxymethylated PGP 9.5 and then resolution by water/acetonitrile gradient from a reverse-phase (PLRP-S300 Å, Polymer Laboratories, England) HPLC column using detection at 220 nm.

Amino acids released from the C-terminus of PGP 9.5 by carboxypeptidase A were identified as their dabsylisothiocyanate derivatives [9].

2.2. *Oligodeoxynucleotide probe synthesis and screening of a cDNA library*

A 'best guess' 35-mer, 5'-GATGGC CGGAT CCATTC CCTGTGAAC CATGGCGC-3', for the sequence DGRMPFPVNHGA was synthesized automatically, purified by electrophoresis and 5'-end labelled [10]. 2×10^5 plaques of a human retinal cDNA library constructed in λgt 10 [11] were screened [10] at a stringency corresponding to 70% probe match [12] and rescreened at higher temperatures. The two best positives were further characterized, one of which provided a fully deduced protein sequence.

2.3. *Analysis of cDNA clones*

Amplification, restriction mapping, subcloning into mp8 and sequencing by dideoxy chain termination were performed by standard methods [10,13]. Homology searches both at protein and DNA level employed the EMBL, Genbank and Doolittle databases and searching programmes currently implemented through the Cambridge University Computer Network.

3. RESULTS

3.1. *Analysis of the protein PGP 9.5*

The amino acid content [6] is very similar to that predicted from the cDNA sequence. The N-terminus appeared to be blocked, and only alanine could be released from the C-terminus by carboxypeptidase A. CNBr treatment gave one favourable cleavage releasing a 5 kDa C-terminal fragment with the sequence PFPV....., which was subsequently encountered within both tryptic and chymotryptic fragments (fig.1) and provided a defined site suitable for oligonucleotide probe deduction. A series of short and long peptides from tryptic and chymotryptic digests correlate (fig.2) with the protein sequence deduced from the cDNA.

3.2. *Cloning and sequencing of cDNA for PGP 9.5*

The labelled 35-mer produced hybridization signals up to 85% probe/target homology for the two best positive plaques. From the initial screening, one positive (9.5 λ2) provided an apparently full-length deduced sequence for a 23.6 kDa protein containing all of the sequenced PGP 9.5 peptides. The analysis of this (and subsequent) cDNA clones was simplified by the presence of two internal *Eco*RI sites which (although hindering excision

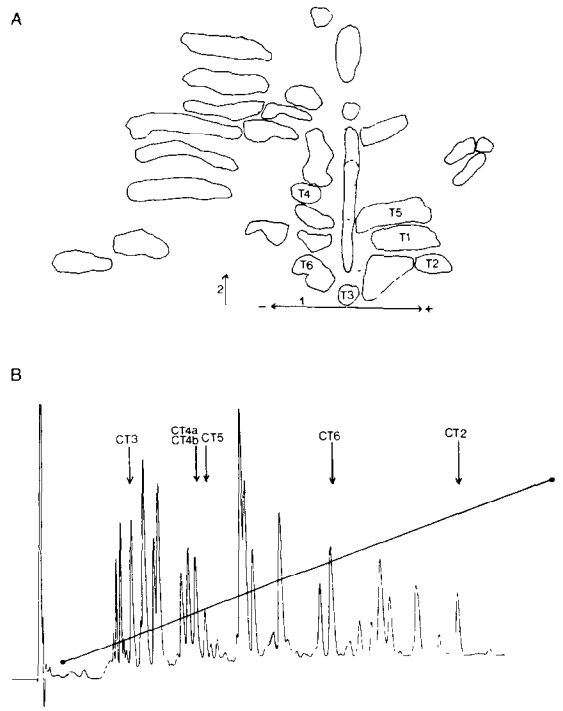

Fig.1. Isolation and sequencing of peptides from PGP 9.5 (see text for details). The numbered fragments were sequenced and are shown in fig.2. (A) Two-dimensional paper map of tryptic peptides. (B) Reverse-phase HPLC profile of chymotryptic peptides. The linear gradient is from 15% to 40% acetonitrile in water/0.1% trifluoroacetic acid.

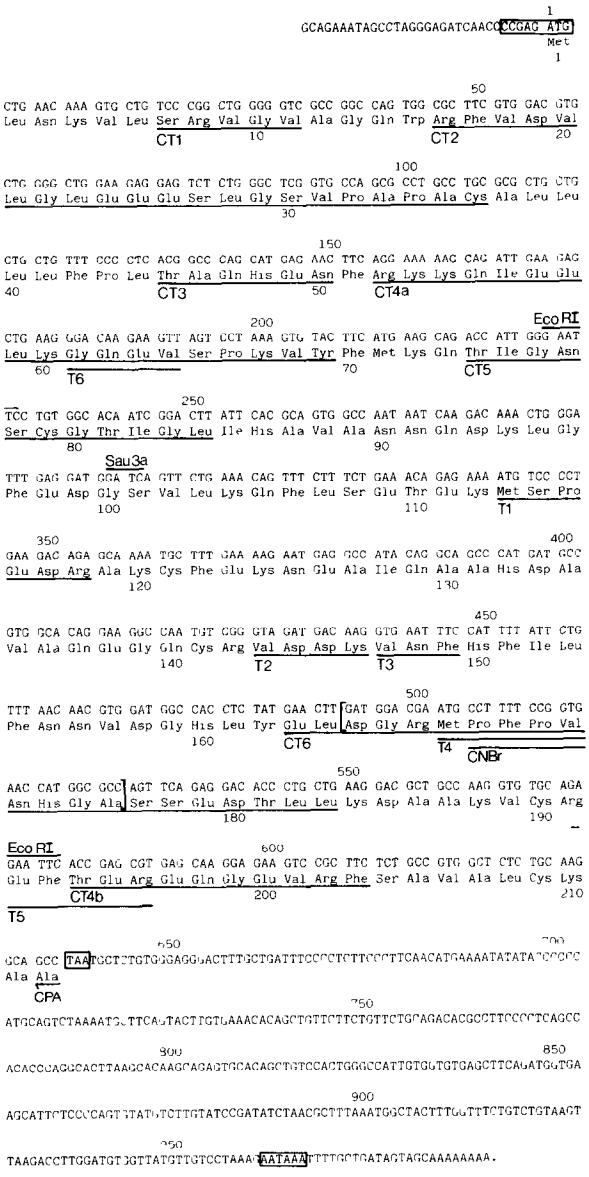

Fig.2. PGP 9.5 cDNA sequence and protein sequence. The oligonucleotide probe was derived from the region in brackets. Two internal *Eco*RI sites and a *Sau*3A site used to establish overlapping subclones for DNA sequencing on both strands are indicated. The Kozak initiation consensus, stop codon and polyadenylation signal (see text) are boxed. The deduced protein sequence is underlined for peptides which have been isolated and sequenced from the protein itself (see fig.1): also indicated are a paper purified chymotryptic peptide (CT1), the C-terminal analysis (CPA), and a cyanogen bromide fragment (CNBr), see text.

of the intact cDNA from the λgt 10 *Eco*RI insertion site) enabled systematic sequencing of the three *Eco*RI fragments on both strands in mp8. Clone 9.5 λ2 did not contain a polyadenylation signal, but rescreening of the cDNA library at high stringency using the internal *Eco*RI fragment of 9.5 λ2 as the probe identified 10 positives out of 10^5 plaques, each containing the same *Eco*RI sites and this internal *Eco*RI fragment. One with a somewhat longer 3′-representation of PGP 9.5 mRNA encoded a 50-mer poly(A) tail preceded by the polyadenylation signal AATAAA and showed that clone 9.5 λ2 extended almost to this signal. The cDNA clones (sequence shown in fig.2) are consistent with an mRNA a little larger than 1 kb with an abundance of about 1/10000 in retinal poly(A)$^+$ mRNA. No significant homologies were identified in the current DNA and protein databases.

4. DISCUSSION

We describe the full-length primary structure of human PGP 9.5 and its mRNA. The protein and cDNA sequences are mutually consistent. The C-terminus has been confirmed on three premises: the presence of a stop codon in the DNA sequence; release only of alanine from PGP 9.5 by carboxypeptidase A (which would not cleave the adjacent lysine); and release of a 5 kDa fragment of sequence PFPV..... from the intact protein by limited cleavage with cyanogen bromide. The N-terminus of PGP 9.5 appears to be blocked and has not been confirmed directly, but the circumstantial data including an in-frame sequenced peptide starting 7 amino acids downstream, an in-frame stop codon 25 nucleotides upstream and DNA sequence at positions −5 to −1 fitting the Kozak initiation consensus [14] substantiate the initiation methionine of the deduced protein sequence. The molecular mass of the predicted protein is 23.6 kDa compared with estimates of 24–27 kDa from denaturing gel electrophoresis of PGP 9.5 [2]. The lack of homology with any previously described protein confirms the identity of PGP 9.5 as a new marker for neurones and cells of the diffuse neuroendocrine system.

The protein is estimated to represent 0.5–1% of total soluble brain protein [2]: here we estimate that its mRNA is above 0.01% of total retinal

mRNA, and since only a proportion of brain and retinal cells contain PGP 9.5 [3,4] the neuronal concentration of both protein and mRNA are likely to be higher.

The primary protein structure opens the possibility for more refined immunochemical studies, and the cDNA probes provide an important tool in studying the developmental biology of PGP 9.5 as a general cytoplasmic marker of neurones and neuroendocrine cells, a diverse group of cells long noted to possess developmental similarities [15], which may now be investigated at the molecular level.

ACKNOWLEDGEMENTS

We thank Dr Jeremy Nathans for the human retinal cDNA library, and Dr Peter Jackson for helpful advice on the protein chemistry. I.N.M.D. is the recipient of an MRC Training Fellowship. The latter part of this work was supported by a grant from the Wellcome Trust.

REFERENCES

[1] Jackson, P.J. and Thompson, R.J. (1981) J. Neurol. Sci. 49, 429–438.
[2] Doran, J.F., Jackson, P., Kynoch, P.A.M. and Thompson, R.J. (1983) J. Neurochem. 40, 1542–1547.
[3] Thompson, R.J., Doran, J.F., Jackson, P., Dhillon, A.P. and Rode, J. (1983) Brain Res. 278, 224–228.
[4] Osborne, N.N. and Neuhoff, V. (1985) Funkt. Biol. Med. 4, 7–13.
[5] Schmechel, D., Marangos, P.J. and Brightman, M. (1978) Nature 276, 834–836.
[6] Thompson, R.J. and Day, I.N.M. in: Neurobiological Research, vol.II (Manangos, P.J. et al. eds) Academic Press, California.
[7] Rode, J., Dhillon, A.P., Doran, J.F., Jackson, P. and Thompson, R.J. (1985) Histopathology 9, 147–158.
[8] Day, I.N.M. and Thompson, R.J. (1986) Biochem. Soc. Trans. 14, 350–351.
[9] Chang, J.Y. (1983) Methods Enzymol. 91, 455–466.
[10] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[11] Nathans, J., Thomas, D. and Hogness, D.S. (1986) Science 232, 193–202.
[12] Lathe, R. (1985) J. Mol. Biol. 183, 1–12.
[13] M13 Cloning and Sequencing Handbook (1984) Amersham International PLC, England.
[14] Kozak, M. (1981) Nucleic Acids Res. 9, 5233–5252.
[15] Pearse, A.G.E. (1968) Proc. Roy. Soc., Ser. B 170, 71–80.