Sequence of a novel carbonic anhydrase-related polypeptide and its exclusive presence in Purkinje cells

Kikuya Kato

MRC Molecular Genetics Unit, Cambridge CB2 2QH, UK

Received 30 July 1990

I isolated a mouse cDNA clone encoding a novel polypeptide which has strong homology with carbonic anhydrase. Unlike the other carbonic anhydrases, it has an additional N-terminal domain with a glutamic acid stretch and an arginine substitutes one of the three histidine residues which bind zinc ion. In the central nervous system, carbonic anhydrase is known to be expressed only in glia cells, but this gene is expressed in neuron, but only in Purkinje cells.

Carbonic anhydrase; Purkinje cell; cDNA cloning

1. INTRODUCTION

The zinc metalloenzyme carbonic anhydrase (CA; carbonate dehydratase, EC 4.2.1.1) catalyzes the reversible hydration of CO_2. Seven isozyme forms (CA I–VII) and several related genes have been identified in mammals [1], and their primary structures show that they belong to a large multiple gene family. The three-dimensional structures of several carbonic anhydrases are known [2,3] and the amino acid residues located in the active site are strongly conserved among different members of the family [1].

In the central nervous system, carbonic anhydrase is expressed only in glia cells, and has been used as a glia cell marker [4]. In particular, extensive developmental studies have been done in retina using it as a marker for Müller cells [5]. Based on immunological, electrophoretic and kinetic criteria, the enzyme in glia cells is a CA II type isozyme [6], and distinct from carbonic anhydrases of the other tissues because over 50% of it is membrane-bound [7].

Mouse cDNA clones have been collected with specific expression patterns in mouse brain [8]. DNA sequence analysis reveals that one of them has strong homology with carbonic anhydrases. However, unlike other carbonic anhydrases, the sequence has an additional N-terminal domain with a glutamic acid stretch. Amino acid residues within the active site are well-conserved, but one histidine residue which binds zinc ion is substituted. Interestingly, this gene is expressed unique in Purkinje neurons of the cerebellum.

Correspondence address: K. Kato, MRC Genetics Unit, Hills Road, Cambridge CB2 2QH, UK

2. MATERIALS AND METHODS

Total RNA was extracted by the guanidinium-CsCl method from mouse brain and poly(A)$^+$ RNA was purified by oligo(dT) cellulose column chromatography [9]. A λgt10 mouse brain cDNA library was constructed using the method of Gubler and Hoffman [10]. To obtain full-length copies, the cDNA was fractionated on a Sepharose CL-4B column and the top 5% was used for the library. The probe used is H30 [8]. Among several positive clones, the longest one was selected and sequenced by the dideoxy chain termination method using the shot-gun strategy [11].

Poly(A)$^+$ RNA (1 μg per lane) was electrophoresed in 1% (w/v) formaldehyde agarose gel, blotted onto a nylon membrane (GeneScreenPlus, Du Pont). Northern hybridization was done by a standard condition [8]. The probe was the same as that used for the library screening.

For in situ hybridization, mouse brain sections were mounted on organosilane-treated microscope slides and kept at −20°C until use. Following fixation and pronase digestion, the sections were prehybridized for 2 h at 37°C in 50% formamide, 5 × SSC, 10 mM Na-PO_4 (pH 7.0), 25 mM DTT, 2 × Denhardt's and 0.1% SDS. Hybridization was performed at 42°C for 30–40 h in the above solution containing 250 μg/ml denatured salmon sperm DNA and ^{35}S-labelled DNA probe (1–1.5 × 10^6 cpm/ml). The probe used was single stranded DNA derived from an M13 clone containing the H30 insert. The sections were washed in three changes of 2 × SSC at room temperature, and then in 50% formamide, 0.5 M NaCl, 10 mM Tris-HCl and 1 mM EDTA (pH 7.4) at 37°C for about 6 h. They were dehydrated in 70% ethanol, and were dipped into Ilford K-5 emulsion and kept at 4°C for 3 weeks. After developing and fixation, sections were stained with Cresyl violet and mounted.

3. RESULTS AND DISCUSSION

In a previous experiment, 35 cDNA clones with specific expression patterns in mouse brain were selected from 950 cDNA clones by Northern and in situ hybridization using the cDNA inserts as probes [8]. One clone, H30, is expressed only in Purkinje cells. The DNA sequence of a full-length clone was determined on both strands. Its longest open reading frame encodes

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

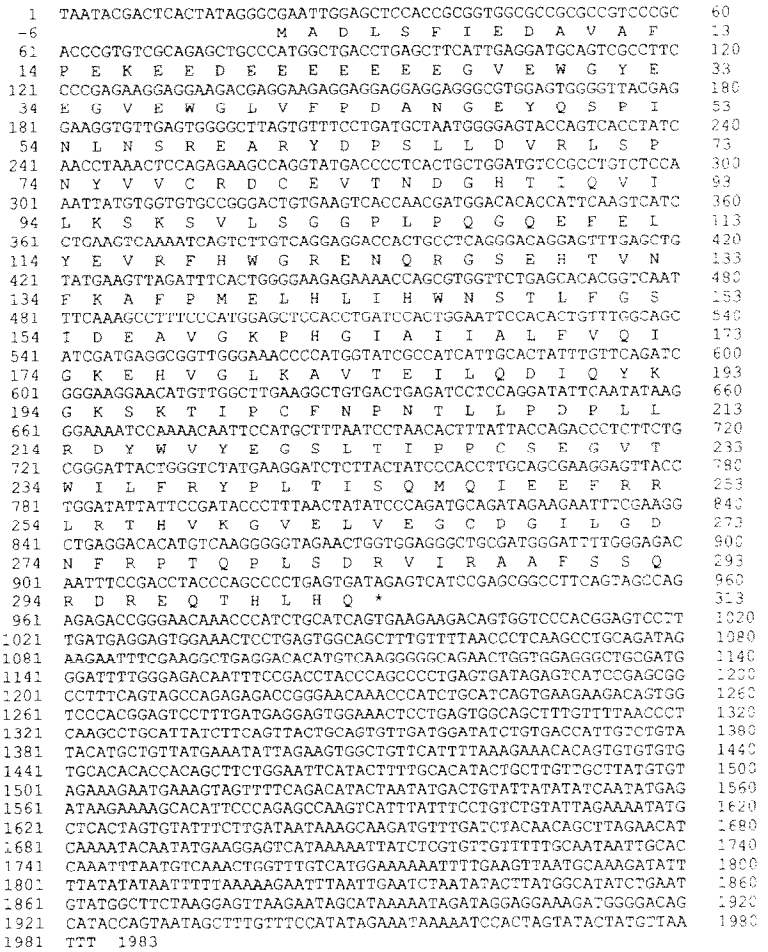

Fig. 1. Nucleotide and predicted amino-acid sequence of a cDNA encoding the carbonic anhydrase-related polypeptide (CARP). The amino acid sequence of the predicted polypeptide is shown as the single letter code above the nucleotide sequence.

303 amino acids (Fig. 1). Comparison of the predicted amino acid sequence with the existing data base reveals that H30 is closely related to carbonic anhydrase. The overall similarity is 37.7% with mouse CA I, 40.7% with mouse CA II, 37.3% with rat CA III and 32.9% with sheep CA VI. This polypeptide is therefore designated as carbonic anhydrase-related polypeptide (CARP). The deduced amino acid sequence and comparison with other carbonic anhydrases are shown in Fig. 2.

It is possible to compare the active site residues of the different CA isozymes. Of 36 such residues at homologous positions, 17 are invariant among all animal sequences [1]. Those of CARP are also well-conserved, but histidine at the position 124 in Fig. 2 and glutamine at the position 122 are substituted by arginine and glutamic acid, respectively. Histidine 124 is one of the three histidine residues which binds zinc ion, and the glutamine residue participates in forming hydrogen-bond network to zinc-bound solvent molecule [1].

Northern hybridization shows that CARP is expressed only in the cerebellum, and is absent from the other parts of the brain (Fig. 3, left). There are two band sizes, of 2.1 and 1.6 kb respectively, possibly reflecting two different polyadenylation sites. In situ hybridization experiment reveals that CARP is localized in Purkinje cells (Fig. 3, right). This makes a clear contrast with carbonic anhydrase, because it is expressed only in glia cells.

The function of CARP is not at all clear, and it is therefore uncertain that CARP even possesses a carbonic anhydrase activity. Because several activities other than the reversible hydration of CO_2 have been

Fig. 2. Comparison of the amino acid sequence of CARP and carbonic anhydrases. Identical residues are indicated by asterisks and conserved residues by dots. The three histidine residues which bind zinc ion are indicated by arrowheads.

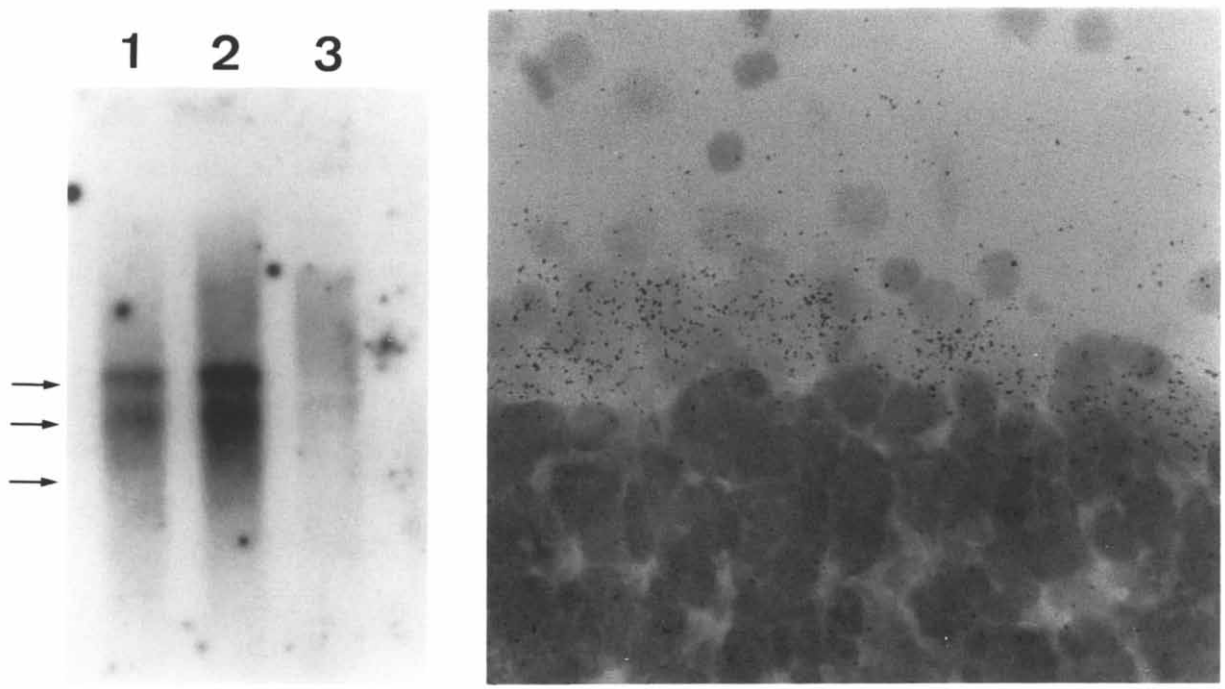

Fig. 3. (Left) Northern blot analysis of CARP. The sources of poly(A)$^+$-RNA are as follows: lane 1, 12-day-old mouse cerebellum; lane 2, 22-day-old mouse cerebellum; lane 3, 22-day-old mouse whole brain minus cerebellum. Left arrows indicate size markers: from top, 1.9, 1.5 and 1.1 kb, respectively. (Right) In situ hybridization analysis of CARP in the mouse cerebellum. Note that silver grains are scattered over only the Purkinje cell layer. The probe was the same as that used for the library screening.

```
                                20                  40                  60
CARP (mouse)      ADLSFIEDAVAFPEKEEDEEEEEEEGVEWGY--EE-GVE-WGLVFPDANGEYQSPINLNS
CAI (mouse)       AS-----------------------ADWGYGSEN-GPDQWSKLYPIANGNNQSPIDIKT
CAII (mouse)      SH-------------------------HWGYSKHN-GPENWHKDFPIANGDRQSPVDIDT
CAIII (rat)       AK-------------------------EWGYGSHN-GPEHWHELYPIAKGDNQSPIELHT
CAVI (sheep)      GH-----------------------GVEWTYSEGMLDEAHWPLEYPKCGGRRQSPIDLQM
                  .                        .*.*     . . *   .*   *  ***....

                               80                  100                 120
CARP (mouse)      REARYDPSLLDVRLSPNYVV-CRDCEVTNDGHTIQVILK---SKSVLSGGPLPQGQEFEL
CAI (mouse)       SEANHDSSL--KPLSISYNP-ATAKEIVNVGHSFHVIFDDSSNQSVLKG-PLADS--YRL
CAII (mouse)      ATAHHDPAL--QPLLISYDK-AASKSIVNNGHSFNVEFDDSQDNAVLKGGPLSDS--YRL
CAIII (rat)       KDIRHDPSL--QPWSVSYDP-GSAKTILNNGKTCRVVFDDTFDRSMLRGGPLSGP--YRL
CAVI (sheep)      KKVQYNPSLRALNLT-GYGLWHGEFPVTNNGHTVQISLPSTMSMTT------SDGTQYLA
                   . .....*       .*     .  . * *.. .. .    . .       .    .

                                  140           160                 180
CARP (mouse)      YEVRFHWGRENQ--RGSEHTVNFKAFPMELHLIHWNSTLFGSIDEAVGKPHGIAIIALFV
CAI (mouse)       TQFHFHWGNSSND-HGSEHTVDGTRYSGELHLVHWNSAKYSSASEAISKADGLAILGVLM
CAII (mouse)      IQFHFHWG-SSDG-QGSEHTVNKKKYAAELHLVHWNT-KYGDFGKAVQQPDGLAVLGYFL
CAIII (rat)       RQFHLHWG-SSDD-HGSEHTVDGVKYAAELHLVHWNP-KYNTSEEALKQPDGIAVVGIFL
CAVI (sheep)      KQMHFHWGGASSEISGSEHTVDGMRYVIEIHVVHYNS-KYNSYEEAQKEPDGLAVLAALV
                   . ..*** ..    ******.   .  *.*..*.*.  ... ..*  ...*.*... ..

                               200                 220                 240
CARP (mouse)      QIGK--EHVGLKAVTEILQDIQYKGKSKTIPCFNPNTLLPDPLLRDYWVYEGSLTIPPCS
CAI (mouse)       KVGP--ANPSLQKVLDALNSVKTKGKRAPFTNFDPSSLLPSSL--DYWTYFGSLTHPPLH
CAII (mouse)      KIGP--ASQGLQKVLEALHSIKTKGKRAAFANFDPCSLLPGNL--DYWTYPGSLTTPPLL
CAIII (rat)       KIGR--EKGEFQILLDALDKIKTKGKEAPFNHFDPSCLFPACR--DYWTYHGSFTTPPCE
CAVI (sheep)      EVKDYTENAYYSKFISHLEDIRYAGQSTVLRGLDIEDMLPGD-LRYYYSYLGSLTTPPCT
                  ..    .        . *....  *.   .  ..   ..*.     *. * **.* **

                               260                 280                 300
CARP (mouse)      EGVTWILFRYPLTISQMQIEEFRRLRTHVKGVELVEGCDGILGDNFRPTQPLSDRVIRAA
CAI (mouse)       ESVTWVICKDSISLSPEQLAQLRGLLSSAEG-----ESAVPVLSNHRPPQPLKGRTVRAS
CAII (mouse)      ECVTWIVLREPITVSSEQMSHFRTLNFNEEG-----DAEEAMVDNWRPAQPLKNRKIKAS
CAIII (rat)       ECIVWLLLKEPMTVSSDQMANVRSLFASAEN-----EPPVPLVGNWRPPQPIKGRVVRAS
CAVI (sheep)      ENVHWFVVADTVKLSKTQVEKLEN--------SLLNHQNKTIQNDYRRTQPLNHRVVEAN
                  * . * .   ....*  *...                    . .. *..**.. * . *.

                               320                 340
CARP (mouse)      FSSQRDRE---QTHLH-----------------------------Q
CAI (mouse)       F-------------------------------------------
CAII (mouse)      F-----------------------------------------K
CAIII (rat)       F-------------------------------------------K
CAVI (sheep)      FMSRPHQEYTLASKLHFYLNNIDQTLEYLRRFIEQKIPKRKKQENWP
                  *
```

reported with CA isozymes [12], the main function of CARP might be different from the reversible hydration of CO_2. Destruction of the CARP gene using some transgenic procedure may be useful to analyze its function [13].

Acknowledgements: The author thanks Dr Sydney Brenner for support and encouragement, and Dr Stephen Hunt for use of equipment. The author is supported by a research fellowship from E.I. Du Pont de Nemours and Co.

REFERENCES

[1] Tashian, R.E. (1990) BioEssays 10, 186–192.
[2] Notstrand, B., Vaara, I. and Kannan, K.K. (1974) in: Isozymes: Molecular Structure, vol. I (Market, C.L. ed.) pp. 579–599, Academic Press, New York.
[3] Deutch, H.F. (1987) Int. J. Biochem. 19, 101–113.
[4] Sapirstein, V.S., Strocchi, P. and Gilbert, J.M. (1984) Ann. NY Acad. Sci. 429, 481–493.
[5] Linser, P. and Moscona, A.A. (1984) Ann. NY Acad. Sci. 429, 430–446.
[6] Sapirstein, V.S., Strocchi, P., Wesolowski, M. and Gilbert, J.M. (1983) J. Neurochem. 40, 1251–1261.
[7] Trachtenberg, M.C. and Sapirstein, V.S. (1980) Neurochem. Res. 5, 573–577.
[8] Kato, K. (1990) Eur. J. Neurosci., in press.
[9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.
[10] Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263–269.
[11] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[12] Tashian, R.E. (1977) in: Isozymes: Current Topics in Biological and Medical Research, vol. 2 (Rattazzi, M.C., Scandalios, J.G. and Whitt, G.S. eds) pp. 31–62, Alan R. Liss, New York.
[13] Rossant, J. (1990) Neuron 2, 323–334.