Two types of complementary DNAs of rat brain protein kinase C

Heterogeneity determined by alternative splicing

Yoshitaka Ono, Tsutomu Kurokawa, Tomoko Fujii, Kenji Kawahara, Koichi Igarashi, Ushio Kikkawa+, Kouji Ogita+ and Yasutomi Nishizuka+

Biotechnology Laboratories, Central Research Division, Takeda Chemical Industries, Ltd, Yodogawa-ku, Osaka 532, and *+Department of Biochemistry, Kobe University, School of Medicine, Chuo-ku, Kobe 650, Japan*

Received 20 August 1986

Two types of complementary DNA clones for rat brain protein kinase C were isolated. These clones encode 671 and 673 amino acid sequences, which differ from each other only in the carboxyl-terminal regions of approx. 50 amino acid residues. This difference seems to result from alternative splicing. Elucidation of the sequences of these cDNA clones as well as some peptides from the purified rat brain enzyme suggests the existence of an additional species of protein kinase C in this tissue. It is attractive to imagine that the heterogeneity of protein kinase C may reflect diverse pathways of signal transduction into the cell.

Protein kinase C *complementary DNA* *Nucleotide sequence* *Alternative splicing*

1. INTRODUCTION

Protein kinase C has attracted great attention in the studies on the activation of cellular functions and proliferation, since this enzyme appears to play key roles in signal transduction for a wide variety of biologically active substances [1–3]. This protein kinase consists of a single polypeptide chain with an approximate M_r of 77 000, and the enzymes from various tissues of animals are apparently similar to one another in their kinetic and catalytic properties [4]. However, the enzyme isolated from rat brain soluble fraction frequently reveals a double band upon SDS-polyacrylamide gel electrophoresis [5]. This raises the possibility that the gene of protein kinase C may be multiple.

In our earlier report [6] partial amino acid sequences of the peptides derived from rat brain protein kinase C were determined, and a cDNA clone encoding the carboxyl-terminal region of this protein kinase was isolated using oligonucleotides deduced from the amino acid sequences as probes. The amino acid sequence predicted from the nucleotide sequence of this cDNA clone revealed a significant homology to other serine/threonine protein kinases [6]. In this paper, we wish to report the complete primary structures of two types of rat brain protein kinase C, that can be predicted from nucleotide sequences of several overlapping cDNA clones, and to discuss that the two types of cDNAs might be derived from the messages which result from alternative splicing.

2. MATERIALS AND METHODS

2.1. *cDNA libraries*

Plasmid cDNA library was constructed using rat brain poly(A)$^+$ RNA by the procedure of Okayama and Berg [7] as described [6]. Phage cDNA library was constructed with λgt10 vector system [8] as follows: the first strand was primed oligo(dT), and the second strand was made with RNase H and DNA polymerase I. The double-stranded cDNA was treated with T_4 DNA

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

polymerase, methylated on internal *Eco*RI sites with *Eco*RI methylase, and ligated to *Eco*RI linkers with T_4 DNA ligase. After double-stranded cDNA was passed through a Bio-gel A50m column, it was cloned in λgt10 phage. Starting from 10 μg of rat brain poly(A)$^+$ RNA, 1.5×10^6 independent phage clones were obtained. Sequential plaque screening was carried out according to Maniatis et al. [9] using each nick-translated restriction endonuclease fragment as probe.

2.2. *Nucleotide sequencing*

Nucleotide sequence analysis was performed using the enzymatic chain termination method in conjugation with M13-derived vectors as described in [6].

3. RESULTS AND DISCUSSION

3.1. cDNA clones

As described in [6], a cDNA clone, which encodes the carboxyl-terminal region of rat brain protein kinase C, has been isolated from the rat brain cDNA library by the colony hybridization procedure with synthetic oligonucleotides as probes. This cDNA clone, designated pCKR9 (fig.1), encodes the carboxyl-terminal region covering about one-third of the entire cDNA for this enzyme. To isolate the cDNAs which may encode the amino-terminal region of protein kinase C, a 0.7 kb *Pst*I fragment of pCKR9 was prepared and employed as probe to rescreen the plasmid library. No cDNA that was longer than pCKR9 was isolated from the library, except for cDNA clones containing an unspliced intron (see below).

Therefore, a new library was constructed with the λgt10 vector system using poly(A)$^+$ RNA that was obtained from rat brain. Upon screening against the λgt10 cDNA library with the 0.7 kb *Pst*I fragment as probe, two types of cDNA (λCKR108 and λCKR107) were isolated. Both cDNA inserts were strongly hybridized to the probe, but the patterns of restriction endonuclease digestion were slightly different. Thus, the two types of cDNA were designated as type I and type II, respectively. Analysis of the nucleotide sequences of these two clones revealed that the λCKR107 insert contained an additional 216 nucleotides to the λCKR108 insert at position 2547 (fig.2). This insertion appeared to be derived by alternative splicing (see below). These cDNA clones both contained a 3′-untranslated region of about 500 nucleotides for type I and 700 nucleotides for type II, and were not long enough to code the entire protein kinase C molecule that is estimated to have an M_r of 77 000 [4].

In order to obtain further the 5′-sequence, the λgt10 cDNA library was rescreened with a ^{32}P-labeled *Eco*RI-*Bgl*II 0.5 kb fragment obtained from λCKR107. Three additional cDNA clones were isolated which extended to the 5′-termini (λCKR152, λCKR166 and λCKR172). Fig.1 shows the series of overlapping clones covering the entire coding region of protein kinase C.

3.2. *cDNA sequences*

The complete nucleotide sequence of rat brain protein kinase C is shown in fig.2. Starting at the

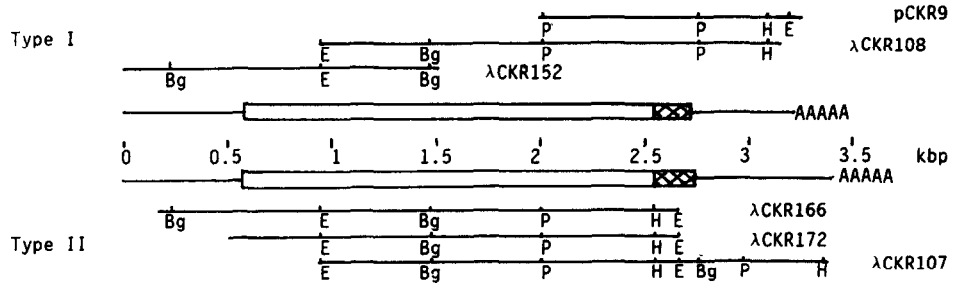

Fig.1. Schematic representation of rat brain protein kinase C cDNA clones. Overlapping clones and schematic diagrams of two types of the complete cDNA structures are given. Thin lines and boxes indicate the untranslated regions and coding regions, respectively. Cross-hatched regions indicate divergent regions between type I and type II of protein kinase C. Several restriction endonuclease cleavage sites are shown on each cDNA clone: E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; Bg, *Bgl*II.

Fig.2. Nucleotide sequence of rat brain protein kinase C type I cDNA and deduced amino acid sequence. The sequence of an additional 216 nucleotides of the type II cDNA clone is also given in the box. The numbers in the parentheses correspond to the type II cDNA. Potential polyadenylation signal is underlined.

5′-end of the cDNA, 684 nucleotides of untranslated sequence are followed by the open reading frame of 671 amino acids for type I and 673 amino acids for type II beginning with a methionine codon. Based on these cDNA sequences, the M_r is estimated to be 76800 for type

I and 76900 for type II. These values match well the value that was previously estimated for the enzyme by sucrose density gradient centrifugation analysis [4]. The coding regions of protein kinase C are followed by 504 nucleotides for type I and 714 nucleotides for type II of 3′-untranslated sequences containing the hexanucleotide ATTAAA (position 3181 for type I and position 3397 for type II) which precedes the site of polyadenylation in most eukaryotic mRNAs.

Four additional ATGs are found in the 5′-untranslated region at positions 309, 356, 396 and 402. However, presumably none of these regions serves as the translational initiation site for the protein kinase, because three stop codons (positions 359, 403 and 501), that are distributed in the three possible reading frames, occur in the 374 nucleotides between these upstream ATGs and the putative translational initiation codon. It is known that the presence of upstream ATGs does not necessarily repress translational initiation at the proper starting site when it is followed by an in-frame termination codon [10]. Thus the ATG at position 685–687 is most likely the initiation codon.

3.3. *Two types of cDNAs*

The sequence of type II is longer than that of type I by an insertion at 2547 (fig.2). In spite of this insertion into the coding region, the amino acid sequence deduced from the nucleotide sequence of type II is only two amino acids longer than that of type I, due to the appearance of a termination codon in the insertion region.

During the course of screening of protein kinase C cDNA, several cDNA clones were isolated which may be derived from a single unspliced molecule. Out of these clones, pCKR41 and λCKR1 were analyzed more precisely. The nucleotide sequences of these two clones were compared with those of type I and type II protein kinase cDNAs. For pCKR41, the nucleotide sequence of the 3′-region from position 2547 in fig.2 is completely identical to that of type I or type II cDNA, but that of the 5′-region from that position is different in each type. The sequence of the divergent point is TTTGGCAG/ which agrees with the consensus sequence of the acceptor site of splicing [11,12]. On the other hand, for λCKR1, the nucleotide sequence of the 5′-region from position 2547 in fig.2 is the same as that of type II cDNA, but that of the 3′-region is different from that of type II cDNA. The sequence of the divergent point is /GTGACA which agrees with the consensus sequence of the splicing donor site [11,12]. These results strongly suggest that the additional sequence in the type II cDNA does not originate from the artificial recombination in the cDNA cloning experiment, but from an alternative splicing event as shown in fig.3a. Analysis of the genomic clones may provide more precise information. The production of different polypeptides from a single gene by the alternative splicing has also been observed in the case of immunoglobulins [13,14], calcitonin [15], preprotachykinin [16], troponin [17], and *c-abl* [18].

The amino acid sequences of type I and type II, which are deduced from the nucleotide sequences of the divergent regions, exhibit a significant homology as shown in fig.3b. Conservation of the amino acid sequence of this region may indicate that this region may indicate that this region is an important domain which is related to the biological activity of protein kinase C. Both of the products of type I and type II cDNAs expressed in mam-

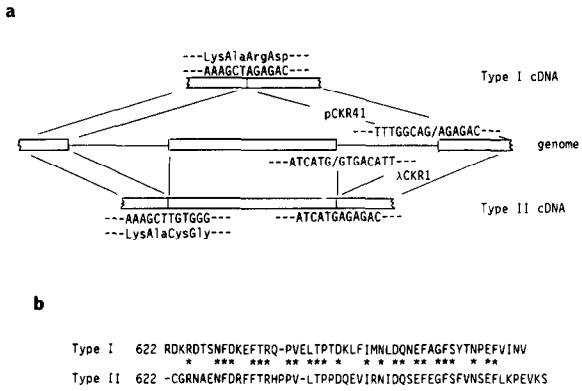

Fig.3. Divergent region between type I and type II of rat brain protein kinase C. (a) Model for the generation of type I and type II mRNA by alternative splicing. Boxes and lines indicate the putative exons and introns, respectively. Partial nucleotide sequences of the cDNA (pCKR41 and λCKR1) are shown, which seem to be derived from the unspliced molecule. (b) Comparison of the amino acid sequences of the divergent regions. Gaps indicated by hyphens are introduced to optimize homology. Stars indicate the amino acid residues which are common for the two sequences.

malian cultured cells showed Ca^{2+}- and phospholipid-dependent protein kinase activity. The precise procedures for the expression of the enzyme will be described elsewhere.

3.4. *Comparison of protein kinase C with other protein kinases*

In most protein kinases, the catalytic domains appear to be located at the carboxyl-terminal regions of the polypeptide chains. Comparing the predicted amino acid sequence of protein kinase C with those of other protein kinases, it is clear that the protein kinase domain of protein kinase C is also located at the carboxyl-terminal region. Fig.4

```
A-kinase    1 GNAAAAKK--GSEQE-SVK--EFL-AKA--KEDFLKKWENPAQNTAHLDQFERIKTL--GT
                **  * *   * *  *   * *      *     * *   * *
C-kinase  303 QKFERAK--IG--QG-T-KAPEEKTANTISKFD--NN-GNRDR--MKLTDFNFLMVL--GK
                *  **    ** *  *   *          ** ***   * *
G-kinase  315 IDRDSFKHLIGGLDDVSNKAYEDAEAKA--KYEAEAA--FFA-N-LKLSDFNIIDTLGVG-

A-kinase   52 GSFGRVMLV--KHME--TGNHYAMKILDKQK-VV-KLKQIEHTLM-E-KR-ILQ-AV--NF
              **** ***   *     *  ** *** * * **      * * * * **  * *
C-kinase  351 GSFGKVMLSERKG----TDELYAVKIL-K-KDVVIQDDDVECT-MVE-KR-VL--ALPGKP
              * ** * *   *     *  * *** * *          *   * *
G-kinase  369 G-FGRVELVQLKSEESKT---FAMKIL-K-K----RH--IVDTRQQEHIRSEKQIMQGAHS

A-kinase  101 PFLVKLEFS-FKDNSNLY-MVMEYVPGGE-MFSHLRRIGRESEPHA-RFYAAQIVLTFEYL
              *** * * *    ** * **** **  *  *    **  ****  **** *     *
C-kinase  401 PFLTQL-HSCFQTMDRLYFM-MEYVNGGDLMY-HIQQVGRFKEPHA-VFYAAEIAIGLFFL
               *  *   *     ** * **   ** *        * * *     ** *        *
G-kinase  418 DFIVRLYRT-FKDSKYLY-MLMEACLGGELWTI-LRDRGSF-EDSTTRFYTACVVEAFAYL

A-kinase  158 HSLDLIYRDLKPENLLIDQQGYIQVTDFGFAKRVKGRTWTL---CGTPEYLAPEIILSKGY
               *  ******  *  * * *   ***  *      *    **** * *****    *
C-kinase  458 QSKGIIYRDLKLDNVMLDSEGHIKIADFGMCKENIWDGVTTKTFCGTPDYIAPEIIAYQPY
               ********** *  **    *  ***  *    *  * ****** * *****
G-kinase  475 HSKGIIYRDLKPENLILDHRGYAKLVDFGFAKKIGF-GKKTWTFCGTPEYVAPEIILNKGH

A-kinase  216 NKAVDWWALGVLIYEMAAGYPPFFADQPIQIYEKIVSGKVRFPSHFSSD-LKDLLRNLLQV
               * ***** *** *** **  **          *   * *  *      *
C-kinase  519 GKSVDWWAFGVLLYEMLAGQAPFEGEDEDELFQSIMEHNVAYPKSMSKE-AVAICKGLMTK
               * * *  * ** * *  ** * *          *              *
G-kinase  535 DISADYWSLGILMYELLTGSPPFSGPDPMKTYNIILRGIDMIEFPKKIAKNAANLIKKLCR

A-kinase  276 DLTK-RFGNLKDGVNDIKNHKWFATTDWIAIYQRKVEA-PFIPKFKGPGDTSNFDDYEEEE
                * * *   *  *** * *  *   **      **    * **      ******
C-kinase  579 HPGK-RLGCGPEGERDIKEHAFFRYIDWEKLE-RKEIQPPYKPKARDKRDTSNFDKEFTRQ
                ***   *  ** *  *   ** *        ** *      ******
G-kinase  596 DNPSERLGNLKNGVKDIQKHKWFEGFNWEGL-RKGTLTPPIIPSVASPTDTSNFDSFPEDN

A-kinase  335 IRVSINEKCGKEFSEF
C-kinase  638 PVELTPTDKLFIMNLDQNEFAGFSYTNPEFVINV
G-kinase  656 DEPPPDDNSGWDID
```

Fig.4. Comparison of the homologous regions among rat protein kinase C type I (C-kinase), catalytic subunit of bovine cyclic AMP-dependent protein kinase (A-kinase) [21], and bovine cyclic GMP-dependent protein kinase (G-kinase) [22]. Gaps indicated by hyphens are introduced to optimize homology. Stars indicate the amino acid residues which are common for these protein kinases.

shows the amino acid sequences within the conserved region of three protein kinases. Large clusters of homology occur throughout the enzymes. The lysine residue for the ATP-binding site is conserved in protein kinase C (amino acid 371), and a characteristic sequence of Gly-X-Gly-X-X-Gly is found 15–20 residues upstream from this binding site. This region is highly conserved in all protein kinases including tyrosine-specific protein kinases and is likely to be an important structural domain [9].

Since protein kinase C is a Ca^{2+}- and phospholipid-dependent enzyme, there may be a Ca^{2+}-binding site in the enzyme molecule. Recently, Geisow [20] has found the consensus sequence of the Ca^{2+}-binding site by comparison of the amino acid sequences of various Ca^{2+}-binding proteins. Searching for such a consensus sequence (Gly-X-Gly-Thr-Asp-Glu) in the predicted protein kinase C structure, an analogous sequence is located at amino acid residues 361–366. However, it may not be concluded at present that this region is indeed the Ca^{2+}-binding site of protein kinase C.

Here, the primary structures of two types of rat brain protein kinase C are presented, which are deduced from the nucleotide sequences of several overlapping cDNA clones. The amino acid sequences of the two types of protein kinase C differ from each other only in the carboxyl-terminal region, and this difference seems to result from alternative splicing. Since alternative splicing events are known to be related to the regulation of tissue specific gene expression, it is suggestive that protein kinase C shows some tissue specificity.

It has been reported [6] that digestion of rat brain protein kinase C with lysylendopeptidase produced a pair of closely similar amino acid sequences, one of which is not coded by any of the cDNA clones described above. It is likely that an additional species of protein kinase C may occur in the brain tissue. Although it remains unclear whether several forms of this enzyme exist in a single cell or are derived from different cell types, it is attractive to imagine that various forms of protein kinase C may transduce different signal messages into the cell.

ACKNOWLEDGEMENTS

We thank Dr S. Kuroda for the nucleotide se-

quencing, Miss K. Itoh for technical assistance, and Mrs S. Nishiyama and Miss M. Inatsugu for secretarial assistance. The research in the Department of Biochemistry, Kobe University School of Medicine has been supported in part by the Ministry of Education, Science and Culture, Japan; Muscular Dystrophy Association; Yamanouchi Foundation for Research on Metabolic Disorders; Central Research Laboratories, Takeda Chemical Industries; Merck Sharp & Dohme Research Laboratories; Ajinomoto Central Research Laboratories; and Meiji Institute of Health Sciences.

REFERENCES

[1] Nishizuka, Y. (1984) Nature 308, 693–698.
[2] Nishizuka, Y. (1986) Science 233, 305–312.
[3] Kikkawa, U. and Nishizuka, Y. (1986) Annu. Rev. Cell Biol. 2, 149–178.
[4] Kikkawa, U., Takai, Y., Minakuchi, R., Inohara, S. and Nishizuka, Y. (1982) J. Biol. Chem. 257, 13341–13348.
[5] Kikkawa, U., Go, M., Koumoto, J. and Nishizuka, Y. (1986) Biochem. Biophys. Res. Commun. 135, 636–643.
[6] Ono, Y., Kurokawa, T., Kawahara, K., Nishimura, O., Marumoto, R., Igarashi, K., Sugino, Y., Kikkawa, U., Ogita, K. and Nishizuka, Y. (1986) FEBS Lett. 203, 111–115.
[7] Okayama, H. and Berg, P. (1983) Mol. Cell. Biol. 3, 280–289.
[8] Watson, C.J. and Jackson, J.F. (1985) in: DNA Cloning, A Practical Approach (Glover, D.M. ed) vol. 1, pp. 79–88, IRL Press, Oxford.
[9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
[10] Liu, C.-C., Simonsen, C.C. and Levinson, A.D. (1984) Nature 309, 82–85.
[11] Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) Proc. Natl. Acad. Sci. USA 75, 4853–4857.
[12] Mount, S.M. (1982) Nucleic Acids Res. 10, 459–472.
[13] Alt, F.W., Bothwell, A.L.M., Knapp, M., Siden, E., Mather, E., Koshland, M. and Baltimore, D. (1980) Cell 20, 293–301.
[14] Maki, R., Roeder, W., Traunecker, A., Sidman, C., Wabel, M., Raschke, W. and Tonegawa, S. (1981) Cell 24, 353–365.
[15] Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S. and Evans, R.M. (1982) Nature 298, 240–244.
[16] Nawa, H., Kotani, H. and Nakanishi, S. (1984) Nature 312, 729–734.
[17] Breitbart, R.E., Nguyen, H.T., Medford, R.M., Destree, A.T., Mahadavi, V. and Nadel-Ginard, B. (1985) Cell 41, 67–82.
[18] Ben-Neriah, Y., Bernards, A., Paskind, M., Daley, G.Q. and Baltimore, D. (1986) Cell 44, 577–586.
[19] Hunter, T. and Cooper, J.A. (1985) Annu. Rev. Biochem. 54, 897–930.
[20] Geisow, M.J. (1986) FEBS Lett. 203, 99–103.
[21] Shoji, S., Parmelee, D.C., Wade, R.D., Kumar, S., Ericsson, L.H., Walsh, K.A., Neurath, H., Long, G.L., Demaille, J.G., Fischer, E.H. and Titani, K. (1981) Proc. Natl. Acad. Sci. USA 78, 848–851.
[22] Takio, K., Wade, R.D., Smith, S.B., Krebs, E.G., Walsh, K.A. and Titani, K. (1984) Biochemistry 23, 4207–4218.