Molecular cloning and the complete nucleotide sequence of cDNA to mRNA for S-100 protein of rat brain

Ryozo Kuwano, Hiroshi Usui, Toshinaga Maeda, Toshikazu Fukui*, Nobuko Yamanari*, Eiko Ohtsuka*, Morio Ikehara* and Yasuo Takahashi

Department of Neuropharmacology, Brain Research Institute, Niigata University, Asahimachi 1, Niigata 951, and * Department of Chemistry, Faculty of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565, Japan

Received 16 August 1984; Accepted 12 September 1984

ABSTRACT

The complete nucleotide sequence of mRNA for β-subunit of rat brain S-100 protein was determined from recombinant cDNA clones. The sequence was composed of 1488 bp which included the 276 bp of the complete coding region, the 120 bp of the 5'-noncoding region and the 1092 bp of the 3'-noncoding region containing two polyadenylation signals. In addition, the poly(A) tail was also found. The amino acid sequence deduced from the nucleotide sequence was homologous to the amino acid sequence of bovine S-100 β subunit except 4 residues showing species differences. From the viewpoint of evolutionary implications, the homology between the nucleotide sequence of S-100 and those of rat intestinal Ca-binding protein (ICAβP) and calmodulin (CaM) was examined. A dot-blot hybridization of poly(A) RNA from the developing rat brains using a labeled cDNA showed a rapid increase in S-100 mRNA at 10-20 postnatal days. The presence of S-100 mRNA in C-6 glioma cells is also described.

INTRODUCTION

S-100 protein is a brain specific protein discovered by Moore et al (1). This protein is mainly localized in astrocytes in the central nervous system, although many investigators are still studying about the problem of cellular localization (2). Recently Isobe et al isolated α and β subunits of S-100 protein from bovine brain and determined the amino acid sequences of each subunit, revealing the structural relation of S-100 with calcium-binding proteins of EF-hand type (3-9). However, rat brain contains exclusively S-100b composing homodimer of β subunit (8, 9).

We previously isolated S-100b from rat brain (9) and then observed the in vitro synthesis of this protein in a reticulocyte lysate cell free system (10).

In this paper, we report the construction, identification and characterization of cDNA clones to S-100 mRNA of rat brain. The complete nucleotide sequence determination of this cDNA is also presented. From the viewpoint of evolutionary implications, the homology between the nucleotide sequence of S-100 and those of other EF-hand type Ca-binding proteins; rat intestinal.
vitamin D-dependent Ca-binding protein (ICaBP) and calmodulin (CaM), is compared. Further, the dot-blot hybridization of mRNA for S-100 from the developing rat brains is described. The presence of S-100 mRNA in the C-6 glioma cells is also described.

MATERIALS AND METHODS

Enzymes and reagents. Reverse transcriptase (RNA-dependent DNA polymerase, EC 2.7.7.49) was purchased from G.W. Beard (Life Sciences). Restriction endonucleases were from Takara Shuzo Co. (Kyoto, Japan). T4 polynucleotide kinase (EC 2.7.1.78), DNA polymerase (EC 2.7.7.7) and terminal deoxynucleotidyl transferase (EC 2.7.7.31) were obtained from Bethesda Research Laboratories (Rockville, MD). S1 nuclease (EC 3.1.30.1) was obtained from Sankyo Co. (Tokyo, Japan). Oligo(dT)-cellulose was type 3 from Collaborative Research (Waltham, MA). [γ-32P]-ATP (5100 Ci/mmol) and [α-32P]-dCTP (3000 Ci/mmol) were obtained from Amersham (U.K.).

Preparation of RNA. Total microsomes were prepared from adult rat brains and microsomal RNA was isolated by phenol-chloroform-isooamylalcohol extraction procedure (11). Poly(A) RNA was isolated from the microsomal RNA by oligo(dT)-cellulose chromatography (12).

Oligodeoxynucleotide synthesis. A mixture of all 16 possible 17-base long oligodeoxynucleotides (3'-AAA GTT CTT AAA TAC CG-5') shown in Fig. 1, one of which is complementary to mRNA for amino acid residues (71-76) of B subunit of bovine S-100, was synthesized by the modified phosphotriester method on polystyrene by using programmed synthesizer (Solid Phase Synthesizer Model 25A, Genetic Design Co.) and purified by using a high performance liquid chromatography (13, 14).

Labeling of oligodeoxynucleotides. The oligodeoxynucleotide was labeled at the 5' end by transfer of [32P] from [γ-32P]-ATP using T4 polynucleotide kinase as described by Wallace et al (15).

Construction and cloning of double-stranded cDNA. Double-stranded cDNA was prepared from poly(A) RNA using reverse transcriptase with oligo(dT) as a primer and inserted into the PstI site of pBR 322, using the dG-dC tailing technique. These procedures were essentially performed as described by Land et al (16). Transformation of Escherichia coli X1776 or HB101 was carried out according to the procedure of Dagert et al (17).

Colony hybridization. In order to carry out colony screening, colony hybridization with [32P]-labeled oligodeoxynucleotide was performed at 39°C.
Fig. 1. Synthetic oligodeoxynucleotides used as probes for screening cloned cDNA for β subunit of S-100. All possible coding sequences and the corresponding 17-base long nucleotide synthesized are given for the carboxy-terminal hexapeptide sequence of β subunit of bovine S-100 (amino acid residue 71-76).

overnight in 6X NET (1X NET = 0.15 M NaCl, 0.001 M EDTA, and 0.015 M Tris-HCl pH 7.5), 1X Denhardt's solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, and 0.02% BSA) and 0.1% SDS and washed at 41 °C with 6X SSC (1X SSC = 0.15 M NaCl, 0.015 M sodium citrate pH 7.2) and 0.1% SDS according to the modified procedure of Wallace et al. (15) and Hanahan and Meselson (18).

Isolation of plasmid DNA. Plasmid DNA was isolated from cloned bacteria by the method of Currier and Nester (19).

Restriction nuclease mapping of plasmid DNA. Conditions for restriction endonuclease cleavage of plasmid DNA were essentially as indicated by the supplier. Fragments were electrophoresed on 1% agarose gels containing ethidium bromide and visualized by UV irradiation.

DNA sequence analysis. The DNA sequences of 5'-[32P] end-labeled appropriate restriction fragments of cloned cDNA were determined by the method of Maxam and Gilbert (20). Sometimes the [32P]-labeled single strand was separated and its sequence was determined.

Northern analysis of brain poly(A) RNA. Rat brain microsomal poly(A) RNA was isolated as described above. Poly(A) RNA was also isolated from the C-6 cells (cultured glioma cells), which produce S-100 protein. Various amounts of poly(A) RNA after formamide treatment were electrophoresed on a 1.3% agarose gel containing formaldehyde, transferred to a nitrocellulose filter and hybridized with the nick-translated [32P] S-100 cDNA insert (pRS-4). After washing, the RNA blot was fluorographed at -80 °C using an intensifying screen (21).

RNA dot-blot hybridization. Total poly(A) RNA isolated by guanidine hydrochloride extraction from the whole brains of the developing rats at various ages or the C-6 glioma cells was spotted directly onto nitrocellulose filters and hybridized with the single-stranded nick-translated [32P] S-100
cDNA according to the procedure of Thomas (22).

In vitro translation. mRNA-dependent reticulocyte cell-free reactions and analysis of translation products were carried out essentially according to the procedures of Masuda et al. (10) and Yoshida et al. (23).

RESULTS AND DISCUSSION

Cloning of S-100 cDNA

Poly(A) RNA isolated from the rat brain microsomes was used to construct a rat brain cDNA library. These procedures were performed essentially according to the method of Land et al. (16). [32P]-labeled 17-base long oligodeoxynucleotides were used to screen the clones containing S-100 cDNA inserts. Two colony-hybridization-positive clones, pRS-1 and pRS-4, were isolated from about 10000 tetracycline resistant transformants. An oligodeoxynucleotide was used as an effective probe in the screening of the cloned cDNA molecules for several peptides and proteins such as endorphin, enkephalin (24), gastrin (25) and nerve growth factor (26). In this experiment, we have also used successfully a 17-base long oligodeoxynucleotide as a probe in the screening of the recombinant cDNA for S-100 β subunit mRNA. Inserted cDNAs were isolated from these cloned colonies and their DNA nucleotide sequences were determined by the procedure of Maxam and Gilbert (20).

Restriction nuclease map and nucleotide sequence determination of plasmid DNA

The size of the inserted cDNA nucleotide sequence was determined by digestion of the plasmid DNA with PstI and electrophoresing the fragments on an agarose gel. Restriction nuclease maps of pRS-1 and pRS-4 were constructed by analyses of their single and double enzyme digests. Since pRS-1 was small, we used pRS-4 having about 400 bp for further analysis. The nucleotide sequence of the insert of pRS-4 was determined from 5'-[32P] end-labeled fragments after digestion with Ddel, BstNI or BstEII as shown in Fig. 2 which shows the strategy of sequencing and the restriction map of pRS-4. After we had determined the nucleotide sequence of this cloned cDNA, we rescreened the cDNA library using the nick-translated inserts. Thirteen colonies showed strong positive reaction to this probe in 80000 transformants. Four long cDNA clones (pRS-11, -39, -49 and -51) contained collectively 1488 bp inserts, which included the 276 bp of the complete coding region, the 120 bp of the 5'-noncoding region and the 1092 bp of the 3'-noncoding region. In addition, the poly(A) tail was found. These indicate evidence for the complete nucleotide sequence considering the size.
Fig. 2. Restriction map and sequencing strategy for cDNA inserts in clones pRS-4, -11, -39, -49 and -51. For the isolation of clone pRS-4 the tetracycline resistant transformants were screened by colony hybridization with synthetic oligodeoxynucleotides. For the isolation of clones pRS-11, -39, -49 and -51, the cDNA library was rescreened by hybridization with the nick-translated PstI fragments. The fragments were labeled at the 5'-site with [γ-32p]ATP and polynucleotide kinase. Their sequences were determined as described in the method. The coding nucleotide sequence for S-100 is indicated by the open box and the synthetic oligodeoxynucleotides used as the probe by the closed box.

of the mRNA as described later. The restriction maps and the nucleotide sequences of these cDNAs are also shown in Figs. 2 and 3. The 276 nucleotide sequence corresponded to the full length nucleotide sequence for coding region of the amino acid sequence of β subunit of S-100 protein, including 17 bases of the synthetic probe, although four amino acids at 8, 63, 79 and 81 deduced from nucleotide sequence were exchanged with the amino acids of bovine protein. At positions 8, 63, 79 and 81, we identified methionine, glutamic acid, serine and valine instead of valine, serine, alanine and isoleucine which had been found in β subunit of bovine S-100 by Isobe et al (6). Although the amino acid sequence of β subunit of rat S-100 has not been determined at present, the differences may be due to the species differences. These results indicate that only four amino acids are different between rat and bovine or porcine proteins, confirming the data about amino acid compositions of these proteins (9), and that S-100 protein is one of the most conservative proteins during evolution. The amino acid sequence of rat S-100 is near to that of human one. We determined 18 amino acid residues of amino terminus of β subunit of rat brain S-100 by the microsequencing method of Hunkapiller and Hood (27) and that sequence was identical to the amino acid sequence deduced from our nucleotide sequence. The codon for serine at position 2 is immediately preceded by
Fig. 3. Nucleotide sequence of rat S-100 mRNA deduced from the cloned cDNA and the predicted amino acid sequence. The numbers below the line indicate positions of amino acid, and the numbers above the line indicate nucleotide positions, beginning with the initial codon. The nucleotides in the 5'-untranslated region are indicated by negative numbers. The amino acids of human, bovine and porcine S-100 different from the predicted rat amino acid sequence are shown. The initiation codon AUG and the termination codons UGA and UAG are underlined. Polyadenylation signals, AAUAAA and AUUAAA are double-underlined. * (pRS-51) and ** (pRS-49) indicate the positions of poly(A) tails. The synthetic oligodeoxynucleotide probe is indicated by the open box.
Table I. Intradomain nucleotide homology between domain I and domain II in rat S-100 cDNA

<table>
<thead>
<tr>
<th>Domains</th>
<th>Percent homology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Helix</td>
</tr>
<tr>
<td>I/II</td>
<td>29</td>
</tr>
</tbody>
</table>

The assignment of domains of S-100 protein is based on the amino acid sequence deduced from the nucleotide sequence of rat S-100 cDNA according to Isobe et al. (6).

AUG coding for methionine. It is possible that the amino terminus of mature S-100 is generated directly by removal of the initiator methionine. Further, there is no additional amino acid sequence at carboxyl terminus of the mature protein.

The recent paper described that in contrast to the high conservative nature of the coding region, the 5'- and 3'-nontranslated regions of the cDNA for chicken and eel CaM (28), a calcium-binding protein of EF hand type, have minimal homology. Although the noncoding regions are long and contain 18-repeated A sequence in S-100 cDNA (in two clones, 12-repeated A sequence), similar to those of CaM, it is not clear at present whether such a situation is found in S-100 cDNA or not, because cDNA clones for S-100 protein of other species were not yet obtained. Further, although the high homology was seen between domains I/III and II/IV of CaM, such high homology was not found in the nucleotide sequences between domains I and II of S-100 (Table I). Furthermore, we examined the homology of nucleotide sequences between chicken CaM and B-subunit of rat S-100. High homology was found between the helixes of CaM IV/S-100 I and between the loops of CaM II/S-100 II (Table II), although high homology was not seen in any combinations of other parts (data not shown).

During the preparation of this manuscript, Desplan et al. (29) published a paper about the nucleotide sequence of cDNA for rat ICaBP which is another calcium-binding protein of EF hand type. Since about 32% homology of the amino acid sequences between bovine ICaBP and bovine S-100 protein was found, we examined the homology of the amino acid sequences and nucleotide sequences between rat ICaBP and B subunit of rat S-100. It was interesting to find 36% homology of amino acid sequences and 52% overall homology of nucleotide sequences, especially high homology of nucleotide sequences in
Table II. Interdomain nucleotide homology between domains of S-100 and domains of the other calcium binding proteins

<table>
<thead>
<tr>
<th>Domains</th>
<th>Percent homology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Helix</td>
</tr>
<tr>
<td>ICaBP I/S-100 I</td>
<td>46</td>
</tr>
<tr>
<td>ICaBP II/S-100 II</td>
<td>62</td>
</tr>
<tr>
<td>CaM IV/S-100 I</td>
<td>50</td>
</tr>
<tr>
<td>CaM II/S-100 II</td>
<td>38</td>
</tr>
</tbody>
</table>

The assignment of domains of S-100, CaM and ICaBP is based on our data as in Table I and on the data of Putkey et al. (28) and Desplan et al. (29), respectively.

Furthermore, we examined the homology of 3'- and 5'-noncoding regions among rat S-100, rat ICaBP and chicken CaM, but could not find any high homology (data not shown), although the size of their 3'-noncoding regions were short (104 bp) in ICaBP and long (857 bp) in chicken CaM. These comparative analysis of the nucleotide sequences may be important from the viewpoint of evolutionary implications.

Size of S-100 mRNA

The nick-translated cDNA was used as a probe to establish the size of S-100 mRNA in the rat brain. Fig. 4 shows that it is about 1500 bases in length by Northern transfer technique. A peptide of 10700 molecular weight requires 276 bases for its coding region. The size of mRNA from the C-6 cells was also similar to that of rat brain (Fig. 4). Therefore, S-100 mRNA must have about 1200 noncoding bases including the 3'-poly(A) tail. Fig. 3 shows the 120 nucleotides in the 5'-noncoding regions containing two stop codons, UGA and UAG, found 15 and 60 bases-upstream to initiating codon AUG, respectively. These results suggest that there is no precursor and signal peptide for S-100 protein unexpectedly. In the 3'-noncoding region, 1092 nucleotides were found including 18-repeated A sequence, two polyadenylating signals, AAUAAA and AUUAAA, as shown in many eukaryotic mRNAs (30). In addition, the poly(A) tail was also found. Two polyadenylation signals may be used for identification of the cleavage site and the addition of poly(A) tails as shown in the eel CaM mRNA (31): the first signal, AAUAAA, and the second signal, AUUAAA are followed by the poly(A) tail at the position 1353 (* in Fig. 3) in the pRS-51, or by the poly(A) tail at the
position 1368 (** in Fig. 3) in the pRS-49. Therefore, these data indicate that our cDNA contains the complete nucleotide sequence.

The developmental changes of S-100 mRNA in rat brain

Since cDNA probe was available, the developmental changes of this mRNA were examined by a dot-blot hybridization of poly(A) RNA from rat whole brain using the labeled cDNA. Fig. 5 shows a rapid increase at 10-20 days in S-100 mRNA concentration. A similar developmental change of translatable S-100 mRNA in a cell-free translation system was found, confirming a rapid increase
of S-100 mRNA level in the young adult (data not shown). It is interesting, without the use of translation procedure, to get similar results using the direct quantitative analysis of mRNA for S-100.

The preparation of cloned cDNA probes specific for β subunit of S-100 will permit us the cloning of genomic DNA for S-100 protein and the elucidation of the molecular mechanism of transcription and specific expression of S-100 gene in the glial cells. However, it is evident from the data of Fig. 4 and the data with the dot-blot hybridization (data not shown) that the expression of S-100 gene is found in C-6 glioma cells.

ACKNOWLEDGEMENTS

We would like to thank Dr. Y. Nabeshima, Department of Biochemistry, Niigata University School of Medicine, and Dr. T. Isobe, Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, for their helpful discussions. This work was supported in part by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Science and Culture to Y.T.

REFERENCES
