Vol. 196, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 15, 1993 Pages 223-232

ANTIGENIC PROTEIN SPECIFIC FOR C3H STRAIN MOUSE IS A MITOCHONDRIAL STRESS-70 PROTEIN*

Yuichi Michikawa, Tadashi Baba, Yuji Arai, Teruyo Sakakura‡, Masashi Tanaka†, and Moriaki Kusakabe‡§

Institute of Applied Biochemistry, University of Tsukuba, Tsukuba Science City, Ibaraki 305, Japan

‡Laboratory of Cell Biology, Institute of Physical and Chemical Research (RIKEN), Tsukuba Science City, Ibaraki 305, Japan

†Department of Biomedical Chemistry, Faculty of Medicine, University of Nagoya, Nagoya 466, Japan

Received September 1, 1993

Cells derived from C3H strain mouse produce an antigenic protein (CSA) specific for this strain [Kusakabe, M., *et al.* (1988) *J. Cell Biol.* 107, 257-265]. To examine the molecular basis of CSA, cDNA cloning of the antigenic protein was carried out. The deduced amino acid sequence demonstrates that CSA is the same protein as peptide-binding protein74 (PBP74), a novel member of the stress-70 family. However, comparison of the PBP74/CSA sequences between C3H/HeN and BALB/c strain mice reveals the substitution of two amino acids in the substrate-recognition domain of the stress-70 protein. Western blot analysis indicates that one out of these two residues, arginine at residue 578 in the PBP74/CSA sequence of C3H mouse, contributes to the immunogenecity of CSA. Moreover, the subcellular localization of PBP74/CSA in mitochondria is also demonstrated by immunohistochemical analysis using anti-CSA monoclonal antibody. Thus, it is interesting that a genetic marker sequence in mice is located on the gene encoding a mitochondrial stress-70 protein.

Chimeric animals are utilized to elucidate cell lineage, homeostasis in tissue architecture, and cell/cell interaction. We have previously obtained a polyclonal antibody by immunization of partially purified proteins from muscle and liver extracts of C3H/HeN strain mice into (BALB/c x SJL/J) F1 mice (1). This polyclonal antibody specifically recognizes cells derived only from C3H strain mouse. Cells of other strain mice, including BALB/c and C57BL/6, show no immunoreactivity with the antibody (1). The antigenic protein (C3H strain specific antigen, termed CSA) is localized in cytoplasm of epithelial and mesenchymal cells in various organs. Thus, the anti-CSA

*This study was partly supported by a grant from the Ministry of Education, Science, and Culture of Japan to M.K. and T.B. Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. D17556.

§To whom all correspondence should be addressed. Fax: +81-298-36-9010.

0006-291X/93 $4.00

antibody is a powerful marker for cell biological analysis in chimeric mice. However, the molecular basis of CSA remains to be elucidated. Does a protein molecule specific for C3H strain mouse exist? Is the sequence specific for C3H strain mouse present in a single protein molecule?

In this paper, we describe that CSA is a mitochondrial protein essentially identical to peptide-binding protein74 (termed PBP74, see Refs. 2-7), a novel member of the stress-70 protein family. Furthermore, the antigenic determinant of CSA is generated by nucleotide substitution of the gene encoding PBP74/CSA.

MATERIALS AND METHODS

Materials. All mouse strains, including C3H/HeN, BALB/cA, and C57BL/6J, were purchased from the CLER Japan Inc. (Tokyo). Restriction nucleases and modifying enzymes were purchased from Nippon Gene (Toyama), Toyobo (Osaka), or Takara Shuzo (Kyoto). Goat anti-mouse IgG antibodies conjugated with horseradish peroxidase or fluorescein isothiocyanate (FITC), and rhodamine-conjugated goat anti-rabbit IgG antibody were purchased from Cappel. Radioisotope, [α-^{32}P]dCTP (3,000 Ci/mmol), was purchased from Bresatec, Australia. Monoclonal antibody against CSA was prepared by injection of the antigen into inguinal lymph nodules of (BALB/c x SJL/J) F1 mice, as described previously (1). Hybridomas were produced by fusing spleen cells from the immunized mouse and myeloma cells, MOPC-21 NS-1, in 50% polyethylene glycol. Cells producing anti-CSA antibody were screened by immunohistochemical staining of cultured primary cells derived from C3H/HeN and of tissue sections of C3H, C57BL/6, and BALB/c strain mice (details for preparation of this monoclonal antibody will be reported elsewhere). The staining patterns of various tissue sections were identical between the polyclonal and monoclonal antibodies; the antibodies recognized cells only from C3H strain mouse. Rabbit polyclonal antibody against bovine ubiquinol-cytochrome *c* oxidoreductase (complex III) was prepared as described previously (8). All other reagents were of the highest purity available.

Double-immunofluorescence labeling of cells. Cultured cells were prepared from 13.5-day embryos of C3H/HeN mice as described previously (1). Cells were cultured overnight in Dulbecco's modified Eagle's medium containing 10% fetal calf serum on Tissue Tek chamber slides (Miles Laboratories, Malvern, IL). Cultured cells were fixed in ice-cold 95% ethanol containing 1% acetic acid for 1 h, dehydrated in 99.5% ethanol for 1 h, and then treated in 90 and 70% ethanol. After washing with phosphate-buffered saline (PBS) containing 0.05% Tween 20, fixed cells were incubated with PBS containing 5% normal goat serum and 1% bovine serum albumin (BSA) for 30 min, and with 100-fold diluted mouse monoclonal anti-CSA antibody and rabbit polyclonal anti-complex III antibody overnight at room temperature. Cells were washed with PBS containing 0.05% Tween 20 (PBST), incubated with FITC-conjugated goat anti-mouse IgG F(ab')$_2$ and rhodamine-conjugated goat anti-rabbit IgG for 45 min at room temperature, and washed again with PBST. The slides were examined by confocal laser microscope, LSM10 (Zeiss, Germany).

Isolation of cDNA clones. Total cellular RNA was extracted from various tissues of mice by the guanidinium isothiocyanate procedure followed by CsCl centrifugation (9). Polyadenylated RNA from kidneys of C3H/HeN or BALB/c mice was selected by oligo(dT)-cellulose (type 7, Pharmacia LKB Biotechnology) column chromatography. After denaturing polyadenylated RNA in 10 mM methylmercury hydroxide for 5 min at room temperature, double-stranded cDNA was synthesized using a commercial kit of Pharmacia LKB Biotechnology, ligated into the *Eco*RI site of λgt11, and then *in vitro* packaged. The kidney cDNA library of C3H mouse in λgt11 was screened using monoclonal antibody against CSA as a probe, as described previously (10). Five clones, TM1, TM2, TM4, TM6, and TM7, were finally obtained and plaque-purified. The cDNA insert of TM7 was labeled with [α-^{32}P]dCTP by the random priming method (11), and used to screen recombinant phage from the BALB/c mouse kidney cDNA library

according to the plaque hybridization method (12). Plaque lifts were prehybridized at 42°C in 5 x SSPE (1 x SSPE = 10 mM sodium phosphate, pH 7.7, 0.18 M NaCl, and 1 mM EDTA), 0.02% Ficoll 400, 0.02% polyvinylpyrrolidone, 0.02% BSA, and 0.1% SDS, followed by hybridization at 60°C overnight in the prehybridization buffer containing denatured salmon testis DNA (0.1 mg/ml) and the ^{32}P-labeled probe (2 ng/ml). The filters were washed in 2 x SSC (1 x SSC = 15 mM sodium citrate, pH 7.0, and 0.15 M NaCl) at room temperature for 10 min, in 2 x SSC containing 0.1% SDS at 60°C for 10 min, and in 2 x SSC at room temperature for 10 min, prior to autoradiography at -80°C. A positive clone, YT1, was obtained, and the cDNA insert was subcloned into pUC19 for further characterization.

Constructions of expression plasmids. Expression plasmids for polypeptides encoded by the cDNA sequence of CSA were constructed in a pMAL-c vector (New England Biolabs, Berverly, MA) that contains the *MalE* gene under the control of the *tac* promoter. The cDNA insert of TM7 isolated from a kidney cDNA library of C3H mouse (YT1 for BALB/c mouse) was digested with *Eco*RI and *Hin*dIII. The *Eco*RI/*Hin*dIII fragment was blunt-ended with klenow fragment. The resulting DNA was inserted into pUC19 at the *Hin*cII site, and then digested with *Sma*I and *Hin*dIII to correct the frame of the protein-coding sequence. The *Sma*I/*Hin*dIII fragment was introduced into the pMAL-c vector at the *Stu*I/*Hin*dIII site for expression in *E. coli* TB1. The chimeric constructs were produced by using *Eco*RI/*Eco*47III and *Eco*47III/*Hin*dIII fragments of TM7 and YT1 (for details see Fig. 4). The resulting plasmids were also introduced into *E. coli* TB1.

Expression of fusion protein and Western blot analysis. A single colony of transformants was cultured at 37°C overnight in Luria/Bertani's broth (LB) containing 50 μg/ml ampicillin (3 ml) with constant shaking. A portion (30 μl) of the bacterial culture was added to a fresh LB containing 50 μg/ml ampicillin (3 ml), and incubated at 37°C for 3 h with shaking. The fusion proteins with maltose-binding protein were induced by addition of 9 μl of 0.1 M isopropyl-β-D-thio-galactopyranoside, and the cell growth was continued for 2 h. Cells were harvested by centrifugation at 3,000 rpm for 5 min, washed with 0.2 ml of 10 mM Tris/HCl, pH 8.0, containing 1 mM EDTA (TE), and suspended in 0.1 ml of TE. The cell suspension was mixed with 50 μl of an SDS sample buffer consisting of 30 mM Tris/HCl, pH 6.8, 3% SDS, 3% 2-mercaptoethanol, 30% glycerol, and 0.03% bromophenol blue, boiled for 3 min, and then centrifuged. Proteins (10 μg) were separated by polyacrylamide gel electrophoresis (PAGE) in the presence of SDS (13). The gels were stained with Coomassie Brilliant Blue R-250 (Sigma). For Western blot analysis, proteins on the gels were transferred onto Immobilon-P PVDF membranes (Millipore). The blots were blocked with 0.1% skim milk, probed with 1 : 500 dilution of anti-CSA monoclonal antibody at room temperature for 1 h, and then incubated with goat anti-mouse IgG horseradish peroxidase conjugate at a dilution of 1 : 5,000. The immunoreactive protein bands were detected using an ECL Western blotting detection kit from Amersham according to the manufacturer's protocol.

Northern blot analysis. Total cellular RNAs (10 μg) were glyoxylated, separated by electrophoresis on 1.2% agarose gels, and transferred to Hybond-N nylon membranes (Amersham). Hybridization was carried out at 60°C according to the manufacturer's protocol. The membranes were washed in 2 x SSC at room temperature, in 0.1 x SSC containing 1% SDS at 60°C, and in 0.1 x SSC at room temperature prior to autoradiography at -80°C.

Analytical procedures. Nucleotide sequence analysis was carried out by the dideoxy chain-termination method (14), using a commercial kit of Sequenase Version 2.0 (US. Biochemicals) or *Bca*Best (Takara Shuzo). Computer-aided analysis of nucleotide and protein sequences was carried out using a GENETYX program (Software Development Co., Tokyo). Protein concentration was measured by the method of Bradford (15).

RESULTS

The subcellular localization of CSA was examined by double-immunofluorescence labeling analysis, using anti-CSA monoclonal antibody and anti-mitochondrial complex

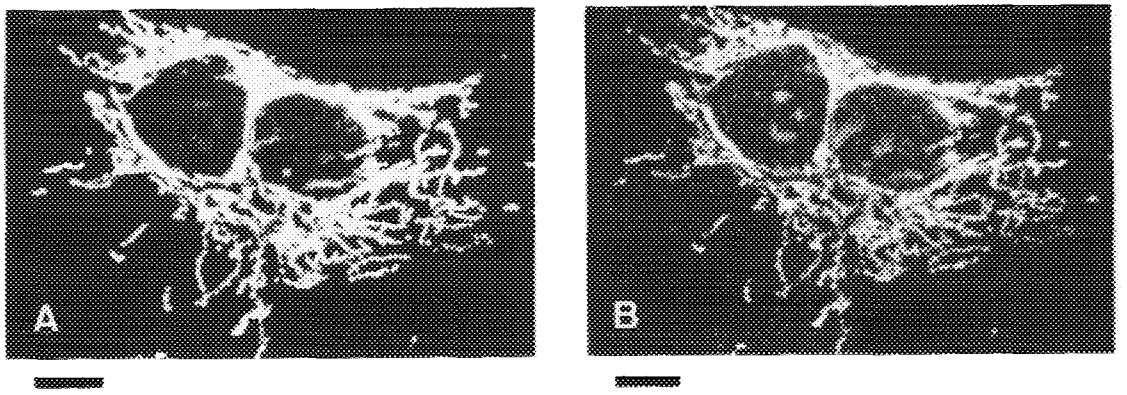

Fig. 1. **Subcellular localization of C3H strain mouse specific antigen (CSA).** Cultured embryonic cells from C3H/HeN strain mouse were double-stained with mouse monoclonal anti-CSA antibody (A) and rabbit polyclonal anti-complex III antibody (B) by using FITC-conjugated goat anti-mouse IgG $F(ab')_2$ and rhodamine-conjugated goat anti-rabbit IgG antibodies, as described under "Experimental Procedures." Bars, 10 μm.

III polyclonal antibody as probes (Fig. 1). Staining of 13.5-day embryos from C3H strain mouse with these two antibodies showed the exactly same pattern. There was no significant staining of nucleus, plasma membrane, and ER/Golgi. Thus, these data demonstrate that CSA is localized in the mitochondria.

To isolate cDNA clones encoding CSA, monoclonal antibody against CSA was used as a probe to screen approximately 2×10^5 recombinant plaques from a C3H-mouse kidney cDNA library in λgt11. Five positive clones, TM1, TM2, TM4, TM6, and TM7, were obtained. Restriction mapping and Southern blot analysis showed that these five cDNA inserts were all related. The nucleotide sequence of the overlapping cDNA inserts contains a 2,037-nucleotide open reading frame, which is flanked by 73-nucleotide 5'-untranslated and 887-nucleotide 3'-untranslated regions (Fig. 2). On the basis of the similarity to the eukaryotic consensus sequence, as described by Kozak (16), the ATG start codon is assigned at nucleotides 74-76. A common polyadenylation signal,

Fig. 2. **Nucleotide sequence of cDNA encoding CSA and the deduced amino acid sequence.** The deduced amino acid sequence is shown below the nucleotide sequence numbered in the 5'- to 3'-direction. Amino acids are numbered from the N-terminus of the mature protein, Ala at position 1 indicated by an underline, and the residues at the N-terminal side from Ala^1 are represented by negative numbers. A sequence of AATAAA identical to the consensus polyadenylation signal is underlined with broken lines. The cDNA inserts of the isolated clones, TM1, TM2, TM4, TM6, and TM7, began at nucleotide 1, 16, 1, 904, and 25, and ended at nucleotide 2,997, 2,796, 2,788, 2,997, and 2,796, respectively. A cDNA clone, YT1, encoding PBP74/CSA from BALB/c mouse contained the sequence from nucleotide 244 to 2,796. Of these clones, only TM7 and YT1 possessed putative poly(A) sequences of 21 and 10 nucleotides, respectively (not shown). Four nucleotides, T, T, G, and C at nucleotides 1,638, 1,648, 1,925, and 1,943 (arrows), in the cDNA sequence of CSA are substituted into C, C, A, and G in the PBP74 sequence (Ref. 7). Alternatively, the cDNA sequence of YT1 from BALB/c mouse is identical to that of TM7, except two nucleotides at positions 1,925 and 1,943. The boxed amino acids, Val and Arg at residues 572 and 578, in the PBP74/CSA sequence from C3H mouse are changed by the nucleotide substitution into Met and Gly in the sequence from BALB/c mouse, respectively.

```
1                                CAC CACCGTGCAC GCAGCTCCGG GCCCGTGGGG TGTTGGTTCT TGCCCTCGTA ACCCCCTCTG TCCAGCCACC   73

74   ATGATAAGCGCCAGCAGAGCCGCGGCCGCGCGTCTCGTGGGCACCGCTGCGTCCCGGAGCCCCGCAGCCGCCCGTCCCCAGGATGGCTGGAATGGCCTT   172
-46   M  I  S  A  S  R  A  A  A  A  A  R  L  V  G  T  A  A  S  R  S  P  A  A  A  R  P  Q  D  G  W  N  G  L   -14

173  AGCCATGAGGCTTTTAGATTTGTTTCAAGAAGAGATTATGCATCAGAAGCAATCAAGGGTGCAGTGGTTGGTATTGATTTGGGTACTACTAACTCCTGT   271
-13   S  H  E  A  F  R  F  V  S  R  R  D  Y  A  S  E  A  I  K  G  A  V  V  G  I  D  L  G  T  T  N  S  C    20

272  GTGGCTGTTATGGAGGGCAAACAAGCAAAGGTCCTGGAGAATGCTGAAGGTGCCAGAACTACCCCTTCTGTGGTTGCCTTTACAGCAGATGGAGAACGA   370
21   V  A  V  M  E  G  K  Q  A  K  V  L  E  N  A  E  G  A  R  T  T  P  S  V  V  A  F  T  A  D  G  E  R    53

371  CTTGTTGGTATGCCAGCAAAACGGCAAGCTGTCACCAATCCAAACAATACCTTCTATGCTACTAAGCGTCTTATTGGACGACGATATGATGACCCTGAA   469
54   L  V  G  M  P  A  K  R  Q  A  V  T  N  P  N  N  T  F  Y  A  T  K  R  L  I  G  R  R  Y  D  D  P  E    86

470  GTACAGAAAGACACTAAGAATGTTCCTTTTAAAATTGTCCGTGCCTCCAATGGTGATGCTTGGGTTGAGGCTCATGGAAAACTCTATTCTCCAAGTCAG   568
87   V  Q  K  D  T  K  N  V  P  F  K  I  V  R  A  S  N  G  D  A  W  V  E  A  H  G  K  L  Y  S  P  S  Q   119

569  ATTGGAGCATTTGTGTTGATGAAGATGAAAGAGACTGCAGAAAATTACTTGGGCCACACAGCAAAAAATGCTGTGATCACAGTCCCTGCTTATTTCAAT   667
120   I  G  A  F  V  L  M  K  M  K  E  T  A  E  N  Y  L  G  H  T  A  K  N  A  V  I  T  V  P  A  Y  F  N   152

668  GATTCACAGCGACAGGCCACTAAGGATGCTGGCCAGATATCTGGGCTAAATGTGCTTCGAGTGATCAATGAGCCTACAGCTGCTGCTCTAGCTTACGGT   766
153   D  S  Q  R  Q  A  T  K  D  A  G  Q  I  S  G  L  N  V  L  R  V  I  N  E  P  T  A  A  A  L  A  Y  G   185

767  CTGGACAAATCTGAAGATAAAGTCATTGCTGTGTATGATTTAGGTGGTGGAACCTTTGACATTTCTATCCTGGAAATTCAGAAAGGAGTGTTTGAGGTG   865
186   L  D  K  S  E  D  K  V  I  A  V  Y  D  L  G  G  G  T  F  D  I  S  I  L  E  I  Q  K  G  V  F  E  V   218

866  AAATCTACCAATGGGGACACTTTCTTAGGAGGGGAAGACTTTGACCAAGCTTTGTTGCGGCACATTGTCAAGGAGTTCAAGAGAGAGACAGGGGTTGAT   964
219   K  S  T  N  G  D  T  F  L  G  G  E  D  F  D  Q  A  L  L  R  H  I  V  K  E  F  K  R  E  T  G  V  D   251

965  TTGACCAAAGACAACATGGCGCTTCAGAGGGTTCGGGAAGCTGCTGAGAAGGCTAAATGTGAACTTTCCTCATCTGTGCAGACTGACATCAACTTGCCA   1063
252   L  T  K  D  N  M  A  L  Q  R  V  R  E  A  A  E  K  A  K  C  E  L  S  S  S  V  Q  T  D  I  N  L  P   284

1064 TACCTTACCATGGATGCTTCTGGACCAAAGCATTTGAATATGAAGCTGACTCGAGCTCAGTTTGAAGGCATTGTCACAGATCTAATCAAGAGAACTATT   1162
285   Y  L  T  M  D  A  S  G  P  K  H  L  N  M  K  L  T  R  A  Q  F  E  G  I  V  T  D  L  I  K  R  T  I   317

1163 GCTCCGTGTCAGAAAGCTATGCAGGATGCAGAAGTCAGCAAGAGTGACATAGGAGAAGTGATTCTGGTTGGTGGCATGACAAGGATGCCCAAGGTTCAG   1261
318   A  P  C  Q  K  A  M  Q  D  A  E  V  S  K  S  D  I  G  E  V  I  L  V  G  G  M  T  R  M  P  K  V  Q   350

1262 CAGACTGTACAAGATCTTTTTGGCAGAGCCCCGAGTAAAGCTGTTAATCCTGATGAGGCTGTAGCCATCGGAGCTGCCATCCAGGGAGGTGTGTTGGCT   1360
351   Q  T  V  Q  D  L  F  G  R  A  P  S  K  A  V  N  P  D  E  A  V  A  I  G  A  A  I  Q  G  G  V  L  A   383

1361 GGTGACGTTACAGACGTGCTGCTCCTGGATGTCACTCCCCTCTCTCTGGGTATTGAGACTCTGGGAGGCGTCTTTACCAAACTTATTAATAGGAACACC   1459
384   G  D  V  T  D  V  L  L  L  D  V  T  P  L  S  L  G  I  E  T  L  G  G  V  F  T  K  L  I  N  R  N  T   416

1460 ACTATTCCAACCAAAAAGAGCCAGGTGTTTTCTACTGCTGCTGATGGACAAACTCAAGTAGAGATTAAAGTGTGTCAGGGGGAACGAGAGATGGCTGGA   1558
417   T  I  P  T  K  K  S  Q  V  F  S  T  A  A  D  G  Q  T  Q  V  E  I  K  V  C  Q  G  E  R  E  M  A  G   449

1559 GACAACAAACTTCTAGGACAGTTCACTTTGATTGGAATTCCCCCAGCCCCTCGTGGAGTGCCCCAGATTGAAGTTACATTTGACATTGATGCCAATGGG   1657
450   D  N  K  L  L  G  Q  F  T  L  I  G  I  P  P  A  P  R  G  V  P  Q  I  E  V  T  F  D  I  D  A  N  G   482

1658 ATTGTGCACGTTTCTGCCAAAGATAAAGGCACTGGTCGTGAGCAACAGATTGTAATCCAGTCTTCTGGTGGATTAAGCAAAGATGATATTGAAAATATG   1756
483   I  V  H  V  S  A  K  D  K  G  T  G  R  E  Q  Q  I  V  I  Q  S  S  G  G  L  S  K  D  D  I  E  N  M   515

1757 GTTAAAAATGCAGAGAAGTACGCTGAGGAAGACCGCAGGAAGAAGGAACGTGTTGAAGCAGTTAATATGGCTGAAGGAATTATTCATGACACAGAAACC   1855
516   V  K  N  A  E  K  Y  A  E  E  D  R  R  K  K  E  R  V  E  A  V  N  M  A  E  G  I  I  H  D  T  E  T   548

1856 AAGATGGAAGAATTTAAGGACCAGTTGCCTGCTGATGAGTGCAACAAGCTAAAGGAAGAGATTTCCAAAGTGAGAGCGCTCCTTGCTCGAAAGGACAGT   1954
549   K  M  E  E  F  K  D  Q  L  P  A  D  E  C  N  K  L  K  E  E  I  S  K  V  R  A  L  L  A  R  K  D  S   581

1955 GAGACAGGAGAGAACATCAGGCAGGCAGCATCTTCCCTACAGCAGGCGTCATTGAAACTCTTCGAAATGGCGTACAAAAAGATGGCATCTGAACGGGAA   2053
582   E  T  G  E  N  I  R  Q  A  A  S  S  L  Q  Q  A  S  L  K  L  F  E  M  A  Y  K  K  M  A  S  E  R  E   614

2054 GGTTCTGGAAGTTCTGGCACTGGGGAACAGAAGGAAGATCAGAAGGAAGAGAAACAGTAA TCGTG GCAGTGCATT GTGGAGCCAG AAGGACATAC   2148
615   G  S  G  S  S  G  T  G  E  Q  K  E  D  Q  K  E  E  K  Q  *                                          633

2149 TATGAAGCTT GGGACTAAAG GGACTTCCTG AGCAGAAAAG GGGCAGACTT CAGTCTTTTT ACTGTATTTT TGCAGTATTC TATATATAAT   2238

2239 TTCCTTAATA TATAAACTTA GTGACAATTG CTAACTCATT TAATGGGTAA TAAAGTCAGC AATAGCAGGT TCATACTGTT CTGTCACTAG   2328

2329 CCTGTTATTT TCAGCTGCAT GTAAAGGGGT GGGATGGGGC TGTGAACCAA TCATTAAGGT AGATTTGGTT TGTGCTGAAA TGGCTGTGAT   2418

2419 TTCAAGGTGG GAAGCCCATT TCACATGCAG TGGAGGTAGT CTGTCATTGA CCTTGAATTG AGATCATATG CAGATGCTTG TTGGCCAAGA   2508

2509 GCACTACTAT AAAGAATGAC CTCTGTATAT TTGCTCCTAC AACTAATGCC TTTAAGACTG AGCTACCTGT ACCATGGTCT GTAGGTGCAG   2598

2599 AAGCTAGGTC AGTGGATAGC AGTTGTGTTA GCCATAGCTT AAAGTATGAT ATGAGAATGA TATAAGCCTC TCATGGGCCT GAGGCATACT   2688

2689 TCTCTAGCCA CCCTCTTGGT TGGCCAATGT CTGGCATCTG TATTCTTGAT GATTGTTCCT TTTTCATCCA CTCTGGATTT TTTAAATAAA   2778

2779 ATTCTGAAAG CCTCTTGATC TCCTTTGTGA ATGGTGATAG CTCAAGGATT ATGACTGCTA TCAGTTTTGT AGGGAGAAAA ATCACTGGCT   2868

2869 AAAAGGTTGA ACAAATGAAA CATGGGGAGT GACTAATAAA ATGCTGGCAT ATATGCTGGA TGTGAAAGTC CACTCAGGAA GCAGTTTGAA   2958

2959 GCCAGGCAGG GCTGTACAGT TAACTCCGTC TTAATAAAA                                                          2997
```

AATAAA, is located at nucleotides 2,287-2,292, 2,773-2,778, 2,903-2,908, and 2,991-2,996. The cDNA-derived sequence indicates that CSA is initially synthesized as a polypeptide of 679 amino acids with a calculated molecular mass of 73,527 Da (Fig. 2).

Surprisingly, the deduced amino acid sequence of CSA is 99.6% identical to the sequence of a new member of the stress-70 protein family, PBP74, that has been recently cloned from cDNA libraries derived from WEHI 231 lymphoma cells (7). Four nucleotides are different in the cDNA sequence between CSA and PBP74; C, C, A, and G at nucleotides 1,638, 1,648, 1,925, and 1,943 in PBP74 are replaced by T, T, G, and C in CSA, respectively (Fig. 2). The substitution of C to T at nucleotide 1,648 in CSA makes no change in the amino acid sequence. The amino acid residues, Ser, Met, and Gly at residues 476, 572, and 578 in PBP74 are changed by the nucleotide substitution into Phe, Val, and Arg in CSA, respectively. Alternatively, the cDNA sequence encoding PBP74/CSA from BALB/c strain mouse is identical to the PBP74 sequence from WEHI 231, except two nucleotides at 1,638 and 1,648 (data not shown, see Fig. 2). WEHI 231 is a cell line induced in a (BALB/c x NZB) mouse (17), both of which are immunologically negative strains for anti-CSA antibody (1). Thus, we conclude that the substitution of two nucleotides at positions 1,925 and 1,943 is present in the cDNA sequence encoding PBP74/CSA between C3H strain and other strain mice lacking the immunoreactivity with anti-CSA antibody.

Immunohistochemical studies, using anti-CSA antibody, showed the ubiquitous localization of this protein in various tissues of C3H strain mouse (1). To ascertain expression of the PBP74/CSA gene in other mouse strains, including the immuno-negative strains of mouse, Northern blot analysis was carried out using the cDNA fragment of TM1 as a hybridization probe (Fig. 3A). A major 3.1-kb signal was found in livers of all mouse strains examined. A minor signal with a size of 2.9 kb was also detectable. Moreover, these two signals were detected in all tissues of C3H mouse when the full length of the TM1 cDNA was used as a probe. These data demonstrate that the PBP74/CSA gene is ubiquitously expressed in the tissues of all mouse strains. However, there is a variation in the gene expression levels among the tissues; the PBP74/CSA gene is more highly expressed in heart and kidney than in testis and brain, and is less expressed in spleen and lung (Fig. 3B). This expression pattern of the PBP74/CSA gene is similar to that of the gene encoding cytochrome *c* oxidase subunit IV (18) localizing in mitochondria. To verify that these two signals were the true messages of the PBP74/CSA gene, the blot was reprobed by the cDNA fragments carrying the 5'- and 3'-end sequences at nucleotides 1-96 and 2,795-2,997, respectively, that shared a low degree of identity with the cDNA sequences encoding other members of the stress-70 protein family (Fig. 3B). The 5'-end sequence hybridized to two mRNA species with the sizes of 3.1 and 2.9 kb as in the case of the full-length probe, whereas only a single 3.1-kb transcript was found when the 3'-end sequence was used as a probe. These data clearly demonstrate that these two mRNAs are the true messages for the PBP74/CSA gene.

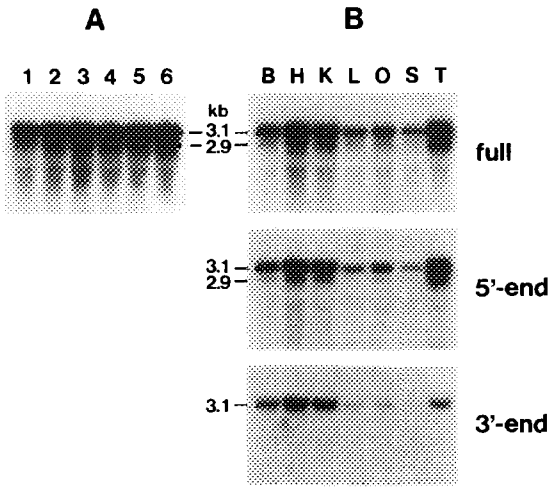

Fig. 3. **Northern blot analysis of total cellular RNAs from various strain mice and distribution of PBP74/CSA mRNA in various tissues of C3H strain mouse.** (A) Total cellular RNAs from livers of C3H/HeN (1), BALB/cA (2), RIIIS/J (3), CE/J (4), DBA/1J (5), and C57BL/6J (6) strain mice were subjected to Northern blot analysis, using the ^{32}P-labeled, entire region of the cDNA insert of TM1 (full) as a probe. Note that C3H/HeN, RIIIS/J, CE/J, and DBA/1J are immunologically positive strains for monoclonal anti-CSA antibody, whereas only BALB/cA and C57BL/6J are negative ones. (B) Total cellular RNAs from brain (B), heart (H), kidney (K), lung (L), ovary (O), spleen (S), and testis (T) of C3H/HeN mice were subjected to Northern blot analysis, using the entire region of the TM1 cDNA (full) as a probe. The blot was reprobed by the cDNA fragments carrying the 5'- and 3'-end sequences at nucleotides 1-96 (5'-end) and 2,795-2,997 (3'-end), respectively (see Fig. 2).

Since two amino acids were different in the amino acid sequence of PBP74/CSA between C3H and BALB/c strain mice (Fig. 2), it was conceivable that either or both of these two residues act as the antigenic determinant for anti-CSA monoclonal antibody. An *Eco*RI/*Hin*dIII fragment encoding the C-terminal 172-residue sequence, stop codon, and 45-nucleotide 3'-untranslated region was introduced into a polylinker site of pMAL-c, and expressed as a fusion protein with maltose-binding protein (Fig. 4). Chimeric constructs with *Eco*RI/*Eco*47III and *Eco*47III/*Hin*dIII fragments in the cDNA sequences of C3H and BALB/c strain mice were also expressed (termed chimera1 and chimera2). Western blot analysis indicated that monoclonal anti-CSA antibody immunoreacted with the fusion proteins from the native C3H and BALB/c-C3H chimeric constructs (chimera2), the latter of which included an Lys-Met-Arg-Ala-Leu-Leu-Ala-Arg-Lys sequence instead of the Lys-Val-Arg-Ala-Leu-Leu-Ala-Arg-Lys sequence at residues 571-579 in the PBP74/CSA sequence of C3H strain mouse (Fig. 2). In contrast to these two constructs, the native BALB/c and C3H-BALB/c chimeric constructs (chimera1) showed a negligibly slight or no immunoreactive band. These results demonstrate that Arg^{578} serves as the key residue in the antigenic sequence for the monoclonal antibody. When the expression plasmids containing other protein-coding regions were expressed in *E. coli*, no immunoreactive signal was found (not shown).

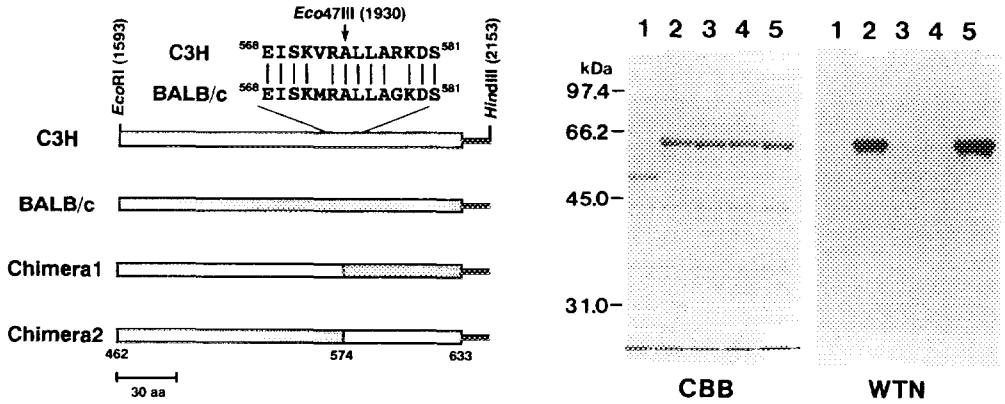

Fig. 4. **Arginine at residue 578 is the key residue in the antigenic sequence for monoclonal anti-CSA antibody.** The *Eco*RI/*Hin*dIII fragments of TM7 (open box) and YT1 (shaded box), which coded for the C-terminal 172-residue sequence, stop codon, and 45-nucleotide 3'-untranslated region (see Fig. 2), were introduced into a polylinker site of pMAL-c and expressed as a fusion protein with maltose-binding protein. Chimeric constructs with *Eco*RI/*Eco*47III and *Eco*47III/*Hin*dIII fragments of TM7 and YT1 were also expressed (chimera1 and chimera2, see left panel). Proteins (10 μg) were subjected to SDS-PAGE (right panel), and the gels were stained with Coomassie Brilliant Blue R-250 (CBB). Moreover, proteins on the gels were transferred onto Immobilon-P PVDF membranes and probed by monoclonal anti-CSA antibody (WTN). 1, pMAL-c vector without insert; 2, C3H; 3, BALB/c; 4, chimera1; 5, chimera2.

DISCUSSION

This study shows that CSA is essentially identical to PBP74 belonging to a member of the stress-70 protein family. PBP74/CSA shares a high degree of sequence identity (65, 60, and 50% identity) with yeast Ssc1p (19), *E. coli* DnaK (20), and mouse hsc70 (21), respectively (data not shown, see Ref. 7). On the basis of immunohistochemical comparison between PBP74 and hsp60 localizing in the matrix of mitochondria (22), PBP74 has been reported to be localized to cytoplasmic vesicles, and possibly not to the nucleus, mitochondria, and plasma membrane (7). An anti-PBP74 antibody shows a vesicular-like staining in contrast to a tubular staining for anti-hsp60 monoclonal antibody (see Fig. 12 in Ref. 7). However, the staining patterns appear to be similar between PBP74 and hsp60, except for the vesicular staining pattern. Our immunohistochemical analysis clearly demonstrates that CSA is localized in the mitochondria (Fig. 1). No vesicular-like staining pattern is found by using anti-CSA monoclonal antibody. Therefore, PBP74/CSA is most likely to be a mitochondrial stress-70 protein. Indeed, the PBP74/CSA precursor possesses a 46-residue leader peptide at the N-terminus (Fig. 2). The 46-residue sequence, which is distinguished from a typical signal peptide sequence by the hydrophobic pattern (data not shown), abundantly contains Ala, Ser, and Arg residues. This fact is consistent with the features of the N-terminal leader sequences of nuclear-encoded, mitochondrial proteins (23, 24). It is thus reasonable to consider that the N-terminal 46-residue sequence is a signal necessary for targeting and transport of the PBP74/CSA precursor into the mitochondria.

The isolated cDNA clones are divided into two groups by the location of the 3'-end; the lengths of the 3'-untranslated regions for TM1 and TM6 are approximately 200 bp longer than those for TM2, TM4, and TM7 (Fig. 2). The 3'-end in TM2 and TM7 is located at nucleotide 2,796 (at nucleotide 2,788 for TM4), and TM7 successively possesses a putative poly(A) sequence (Fig. 2). Since two mRNA species of PBP74/CSA with the sizes of 3.1 and 2.9 kb are found in total cellular RNA (Fig. 3), two polyadenylation signals at nucleotides 2,773-2,778 and 2,991-2,996 may function to generate two transcripts for the gene encoding PBP74/CSA.

The transcript for the PBP74/CSA gene is present in all strains of mouse tested (Fig. 3), including the strains missing the immunoreactivity with anti-CSA antibody (1). This discrepancy is explained by the lack of the antigenic determinant in the PBP74/CSA sequences of the immunologically negative strains for anti-CSA antibody. In fact, Western blot analysis reveals that Arg^{578} is the key residue as the antigenic determinant for monoclonal antibody against CSA (Fig. 4). The amino acid sequences of the stress-70 proteins are functionally divided into two domains (reviewed in Ref. 25): an ATPase domain and substrate-recognition domain in the N- and C-terminal regions of the proteins, respectively. In the present study, the substitution of two amino acids in the PBP74/CSA sequence occurs in the substrate-recognition domain. It is thus an intriguing question whether this modification affects the function of PBP74/CSA as the stress-70 protein. The prediction of the secondary protein structure in the antigenic region, according to the method of Chou and Fasman (26), reveals that the substitution of Arg to Gly at residue 578 destroys an α-helix structure and gives a β-turn (data not shown). However, the significance of the structural change predicted remains to be clarified.

Although several mitochondrial stress proteins, grp75, hsp58, and P71, have already been identified in mammalian tissues and cultured cells (27, 28), the functions of the mitochondrial proteins are not fully understood. Moreover, the identity between PBP74/CSA and these three proteins is not certain at the present time. At any rate, it is interesting that CSA, which was originally isolated as a genetic marker in mice (1), is a mitochondrial stress-70 protein.

Acknowledgment: We thank Kouichi Mizuno for his technical assistance.

REFERENCES

1. Kusakabe, M., Yokoyama, M., Sakakura, T., Nomura, T., Hosick, H. L., and Nishizuka, Y. (1988) J. Cell Biol. 107, 257-265
2. Lakey, E. K., Margoliash, E., Pierce, S. K. (1987) Proc. Natl. Acad. Sci. USA 84, 1659-1663
3. Vanbuskirk, A., Crump, B. L., Margoliash, E., and Pierce, S. K. (1989) J. Exp. Med. 170, 1799-1809
4. Guagliardi, L. E., Koppelman, B., Blum, J. S., Marks, M. S., Cresswell, P., and Brodsky, F. M. (1990) Nature 343, 133-139
5. Van Buskirk, A. M., DeNagel, D. C., Guagliardi, L. E., Brodsky, F. M., and Pierce, S. K. (1991) J. Immunol. 146, 500-506
6. DeNagel, D. C., and Pierce, S. K. (1992) Immunol. Today 13, 86-89

7. Domanico, S. Z., DeNagel, D. C., Dahlseid, J. N., Green, J. M., and Pierce, S. K. (1993) Mol. Cell. Biol. 13, 3598-3610
8. Tanaka, M., Miyabayashi, S., Nishikimi, M., Suzuki, H., Shimomura, Y., Ito, K., Narisawa, K., Tada, K., and Ozawa, T. (1988) Pediatr. Res. 24, 447-454
9. Chirwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) Biochemistry, 18, 5294-5299
10. Huynh, T. V., Young, R. A., and Davis, R. W. (1985) in DNA Cloning (Glover, D. M., ed) Vol. 1, pp. 49-78, IRL Press, Oxford
11. Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13
12. Benton, W. D., and Davis, R. W. (1977) Science 196, 180-182
13. Laemmli, U. K. (1970) Nature 272, 680-685
14. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA. 74, 5463-5467
15. Bradford, M. M. (1976) Anal. Biochem. 72, 248-254
16. Kozak, M. (1981) Nucleic Acids Res. 9, 5233-5252
17. Ralph, P. (1979) Immunol. Rev. 48, 107-121
18. Virbasius, J. V., and Scarpulla, R. C. (1990) Nucleic Acids Res. 18, 6581-6586
19. Craig, E. A., Kramer, J., Shilling, J., Werner-Washburne, M., Holmes, S., Kosic-Smithers, J., and Nicolet, C. M. (1989) Mol. Cell. Biol. 9, 3000-3008
20. Bardwell, J., and Craig, E. A. (1984) Proc. Natl. Acad. Sci. USA 81, 848-852
21. Giebel, L. B., Dworniczak, B. P., and Bautz, E. K. F. (1988) Dev. Biol. 125, 200-207
22. Jindal, S., Dudani, A. K., Singh, B., Harley, C. B., and Gupta, R. S. (1989) Mol. Cell. Biol. 9, 2279-2283
23. Pfanner, N., and Neupert, W. (1990) Annu. Rev. Biochem. 59, 331-353
24. Glick, B. S., Beasley, E. M., and Schatz, G. (1992) TIBS 17, 453-459
25. Gething, M. -J., and Sambrook, J. (1992) Nature 355, 33-45
26. Chou, P. Y., and Fasman, G. D. (1978) Adv. Enzymol. 47, 45-148
27. Mizzen, L. A., Chang, C., Garrels, J. I., and Welch, W. J. (1989) J. Biol. Chem. 264, 20664-20675
28. Leustek, T., Dalie, B., Amir-Shapira, D., Brot, N., and Weissbach, H. (1989) Proc. Natl. Acad. Sci. USA 86, 7805-7808