Molecular Characterization of Transcription Factors That Bind to the cAMP Responsive Region of the Substance P Precursor Gene

cDNA CLONING OF A NOVEL C/EBP-RELATED FACTOR*

(Received for publication, April 17, 1991)

Ryoichiro Kageyama‡, Yoshiki Sasai, and Shigetada Nakanishi
From the Institute for Immunology, Kyoto University Faculty of Medicine, Kyoto 606, Japan

A cAMP response element (CRE) plays an important role in the cAMP-mediated gene regulation. Several factors that recognize a CRE have been characterized, and it has been shown that they need either covalent modification by protein kinase A or a cofactor such as the adeno virus Ela to function as an activator. In this study we show that the substance P precursor gene expression is regulated by protein kinase A and identify the CRE sequence in its promoter region. We find that a novel factor and ATF2 bind to the region containing the CRE of the substance P precursor gene. The sequence analysis indicates that the novel protein, designated CELF, has a significant homology to C/EBP family proteins in the carboxyl-terminal part containing the basic region and the leucine zipper motif. Ubiquitous expression of CELF suggests that this factor is utilized by various genes. Cell-free transcription analyses indicate that CELF is a constitutive transcriptional activator without apparent phosphorylation by protein kinase A. These results demonstrate that multiple factors are responsible for transcriptional control of the substance P precursor gene through the CRE region.

Substance P, a member of the tachykinin neuropeptides, induces depolarization on neurons by binding to its specific receptor (Yokota et al., 1989) and may function as a neurotransmitter in the nervous system (Otsuka et al., 1972; for review see Pernow (1983)). The expression of substance P is controlled by various factors such as nerve growth factor (NGF) and leukemia inhibitory factor (Lindsay and Harman, 1989; Nawa and Patterson, 1990). Substance P and another tachykinin neuropeptide, substance K, are both encoded by the same precursor gene, the preprotachykinin (PPT) A gene (Nawa et al., 1983; 1984). We previously cloned the PPT-A gene and identified the promoter region. Although the precise transcriptional mechanism of the PPT-A gene has not been elucidated yet, we found that the PPT-A gene promoter has several regulatory motifs including a possible cAMP response element (CRE). In this study, to further analyze the regulatory mechanism of the PPT-A gene, we first tested cAMP responsiveness and identified the CRE of the PPT-A gene.

The CRE sequence is present in various genes such as the somatostatin and c-fos genes and mediates CAMP-induced gene expression (Montmyn and Bilezikjian, 1987; Yamamoto et al., 1988). The consensus sequence of the CRE is 5'-TGACGTCA-3', and it functions in a distance- and orientation-independent manner. Several factors that specifically recognize the CRE have been identified. These factors have a leucine zipper motif and form homodimers and heterodimers (Hoeffer et al., 1988; Gonzalez et al., 1989; Maekawa et al., 1989; Hai et al., 1989). Two of them have been further characterized for the transcription activity; the unphosphorylated form of CREB is inactive or acts as a repressor but stimulates transcription of several genes when phosphorylated by protein kinase A (Gonzalez and Montmyn, 1989; Lamph et al., 1990). The other CRE-binding protein, ATF2 (CRE-BP1), is non-functional by itself but can support adeno virus Ela-mediated transcription activation (Liu and Green, 1990). Therefore, both of the factors need either a covalent modification or a cofactor to show transcription activity.

Another leucine zipper motif factor, C/EBP homologues and enhancer binding protein (C/EBP), was recently suggested to be involved in cAMP responsiveness (McKnight et al., 1989). C/EBP is a transcriptional activator which binds to C/EBP homologies and enhancer core sequences and is abundant in hepatocytes and adipocytes (Landschulz et al., 1988a). Three other factors, IL-6DBP (LAP), NF-IL6, and Ig/EBP-1, have also been cloned and found to have a significant homology to C/EBP in the basic domain and leucine zipper regions (Poli et al., 1990; Akira et al., 1990; Descombes et al., 1990; Roman et al., 1990). These proteins have quite similar DNA binding specificities.

Here we report the molecular cloning and characterization of an ubiquitous protein, designated CELF, that belongs to the C/EBP gene family. Our data also show that CELF constitutively activates transcription of the PPT-A gene through the CRE region, raising the possibility of a link between the CRE and the C/EBP gene family members.

MATERIALS AND METHODS

Cloning and Sequence Analyses of CELF and ATF2 cDNAs—cDNA was synthesized by oligo(dT) priming of poly(A)+ RNA of rat whole brain. Double-stranded cDNA was next constructed, ligated to the EcoRI adaptor, size-fractionated to select inserts longer than 1.5 kb, and cloned into the Aqgl11 vector (Promega) (Huynh et al., 1985). Screening was carried out according to Vinson et al. (1988), a modified method of Singh et al. (1988). In brief, Y1090 was infected by recombinant λ phages, plated, and incubated at 42 °C for 4 h. Nitrocellulose filters pretreated in 10 mM isopropyl-1-thio-β-D-galactopy-
ranscripts were overlaid on the plates and incubated at 37 °C for >6 h. The filters were next immersed in binding buffer (10 mM Tris-Cl, pH 7.5, 50 mM NaCl, 3 mM MgCl2, 1 mM dithiothreitol) containing 6 M guanidine hydrochloride for 5 min at room temperature. This solution was diluted with the equal volume of the binding buffer, and then the filters were incubated in this solution for 5 min. The dilution step was repeated four times, and then the filters were transferred to blocking buffer (5% nonfat dry milk in the binding buffer) and left for 30 min. After washing with the filters, the labeled probe was incubated in the same buffer for >2 h. We made the probe containing the CRE of the PPT-A gene as follows. After annealing the oligonucleotide probes, 5'-AGAGACCCGATTAATGGTCAATTTGCGGTCACC-3' and 5'-GGACACATATTGGGTTCGAAAATGACGCAATTATG-3', the second strand was synthesized with Klenow fragment in the presence of 0.5 mM α-32PdCTP (the specific activity of the probe exceeded 5 × 10⁸ cpm/μg) as described before (Kageyama et al., 1989). The filters were then washed with blocking buffer for 1 h and subjected to autoradiography. From 5 × 10⁶ plaques, two positive clones, λ1 and λ2, were obtained. For sequencing analysis we used a chain termination method (Sanger et al., 1977).

**Plasmids**—For pPPTcat, the ~1.8-kb BamHI-Smal fragment containing the bovine PPT-A gene promoter was isolated from λSP15 (Nawa et al., 1984) and cloned into the HindIII site of pRSVcat containing HindIII linkers. We made pA431-PPTcat as follows. Two plasmids were used containing oligonucleotides on either end of the DNA segment, 5'-AGAGACCCGATTAATGGTCAATTTGCGGTCACC-3' and 5'-GGACACATATTGGGTTCGAAAATGACGCAATTATG-3', were annealed and cloned into the BamHI site of pBluescript SK (-) (Stratagene). The BamHI NotI insert was isolated, and two copies of the inserts were cloned into the BglII site of pA431. The expression vector of the catalytic subunit of protein kinase A was kindly provided by Dr. M. R. Montminy (The Salk Institute). pGEMEX1/CELF and pGEMEX1/ATFP contained the EcoRI inserts of X2 and X1 in the pGEMEX1 vector, respectively.

**Preparation and Fractionation of Brain and Liver Nuclear Extracts**—Preparation of brain and liver nuclear extracts and their fractionation by heparin-agarose were done as described (Kageyama et al., 1988). The buffer used was HM, which consists of 20 mM Hepes, pH 7.9, 100 mM KC1, 5 mM MgCl2, 1 mM dithiothreitol, and 8% (v/v) glycerol. The reaction mixture was incubated at 37 °C for 10 min in the presence of at least 4 independent experiments.

**DNA Binding Analysis**—The fusion proteins were prepared as follows. JM109 (DE3) cells transformed by the expression plasmids were grown until OD600 = 0.5-0.7, treated with 1 mM isopropyl-1-thio-β-D-galactopyranoside at 37 °C for 2-3 h, and collected by centrifugation at 4 °C. The cell pellets were suspended in 0.02 volume of 30 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 20% (w/v) sucrose, and incubated on ice for 10 min. After the crude extracts were applied to SDS-polyacrylamide gel electrophoresis, the protein bands were cut out, crushed, and eluted in 5 volumes of the solution containing 50 mM Tris-HCl, pH 7.9, 0.1 mM EDTA, 0.1% SDS, and 150 mM NaCl by incubating at 37 °C for >8 h. The supernatant was recovered and precipitated with 4 volumes of cold acetone. The pellets were dissolved in 0.1 volume of 6 M guanidine hydrochloride, 0.1 M KCl/HM, kept at room temperature for 20 min, and dialyzed against 0.1 M KCl/HM at 4 °C for 8 h.

For the DNase I footprinting experiments, a HindIII-PmacFI fragment of the PPT-A gene promoter located at the HindIII site was used as a probe. Footprinting reactions were carried out in a total volume of 50 μl containing 5 ng of the probe and 50 or 100 ng of fusion proteins, as described before (Dyman and Tjian, 1983).

**Transcription Analysis**—For CAT analysis, a CAT reporter plasmid and expression vectors were cotransfected into CVI cells using the calcium phosphate coprecipitation method, and CAT activities were determined, as previously described (Gorman et al., 1982).

For cell-free transcription analysis, 0.5 or 1.5 μg of a supercoiled DNA template was incubated in a total volume of 25 μl containing 50 μg of A431 nuclear extracts, as described before (Kageyama et al., 1988). After incubation for 60 min at 30 °C, RNA was prepared and hybridized with 5'-end-labeled primer, 5'-TGGACCTGGGATATTCAACCGTGTTG-3'. The primer extension products were analyzed on a 5% polyacrylamide sequencing gel. The phosphorylation reaction with the catalytic subunit of protein kinase A (Sigma, P-2645) was done according to Yamamoto et al. (1988).

**Northern Blot Analysis**—10 μg of poly(A)+ RNA was electrophoresed on a formaldehyde, 1% agarose gel and transferred to a nylon membrane filter. The EcoRI fragment of λ2 was used as a probe.

**RESULTS**

**Analysis of the CRE of the PPT-A Gene**—The sequence examination of the PPT-A gene promoter region suggests that it has a CRE-like sequence (TGCGTCA) between nucleotide residues -195 and -189 relative to the transcription initiation site (Fig. 1A). This sequence is identical to the CRE of the proenkephalin gene (Montminy et al., 1990). We therefore tested the cAMP responsiveness of the PPT-A gene. We carried out DNA-mediated gene transfer experiments using a reporter plasmid containing the CAT gene under the control of the PPT-A gene promoter (pPPTcat). When cotransfected with the catalytic subunit of protein kinase A expression vector, we observed about 4-fold induction of the CAT activity, suggesting that the PPT-A gene is responsive to cAMP.
(Fig. 1B). As a control we used a CAT vector with the enhancerless SV40 early promoter (pAlocat), which does not contain CRE sequences, and we did not see any significant induction by protein kinase A (Fig. 1B). However, when we tested a CAT plasmid containing four copies of the CRE regions of the PPT-A gene in front of the enhancerless SV40 early promoter (pAloc-PPTcat), we observed about 2-fold stimulation by protein kinase A (Fig. 1B). These results therefore suggest that the PPT region containing the sequence TGCGTCA is responsible for the cAMP inducibility, although other regions might be necessary for the full responsiveness because the element we identified reacted weakly to protein kinase A on the heterologous promoter.

To analyze the factors that recognize the CRE of the PPT-A gene, we made nuclear extracts from the whole brain and performed the ultraviolet (UV) cross-linking experiment with the oligonucleotide probe containing the CRE of the PPT-A gene. This probe was substituted with 5-bromo-2'-deoxyuridine and uniformly labeled with [α-32P]dCTP. The brain nuclear extract was mixed with this probe, and cross-linked by UV light. As shown in Fig. 2, three major proteins of 86, 43, and 40 kDa, were labeled by this probe. This labeling was specific because it was competed not by a nonspecific DNA competitor (lane 5) but by a specific fragment containing the CRE (lane 6). When we performed the UV cross-linking experiment with the liver nuclear extract, we detected the 43- and 40-kDa proteins, indicating that these two proteins are ubiquitously expressed. However, we could not detect the 86-kDa protein in the liver nuclear extract, although we observed a band with a somewhat smaller molecular size (lanes 7-10). We next cross-linked the brain nuclear factors which had been partially purified by the sequence-specific oligonucleotide affinity column and found that the 86- and 43-kDa factors were preferentially enriched (lane 1). We also detected a 29-kDa protein after partial purification, but lost the 40-kDa protein. We speculate that the amount of the 29-kDa protein, which could not be detected in the crude brain and liver nuclear extracts, was very low, but it had a high affinity to the CRE. In contrast, the 40-kDa protein probably had a lower affinity to this element. The data shown above thus suggest that at least three brain factors, the 86-, 43-, and 29-kDa proteins, can recognize the CRE region of the PPT-A gene with a high affinity.

Isolation and Characterization of cDNA Clones for CRE Region-binding Proteins—To investigate the relationship between the multiple factors and the CRE region of the PPT-A gene, we screened a λgt11 rat brain expression library using a double-stranded oligonucleotide probe containing the CRE. Two positive clones, λ1 and λ2, were obtained, and subsequently purified and sequenced. As shown in Fig. 3, λ1 encoded a protein of 389 amino acids. We found that the sequence of the λ1 insert was very similar (about 98% identity) to those of human ATF2 (CRE-BP1) and mouse XBP, both of which are CRE-binding proteins (Hai et al., 1989; Maekawa et al., 1989; Ivashkiv et al., 1990). We therefore concluded that this is the rat counterpart of ATF2. Rat ATF2 does not contain the region between amino acid residues 150 and 247 of human ATF2, and the calculated molecular mass of rat ATF2 is 42.3 kDa. The finding that the CRE-binding protein recognizes the PPT-A sequence we used for screening further supports the notion that this sequence is responsible for cAMP induction of PPT-A gene expression.

The second clone, λ2, had an insert of 2141 base pairs, and encoded a protein containing in-frame from the β-galactosidase gene. The first ATG codon present in this frame, located at nucleotide residue 2, was tentatively assigned as the initiation codon; the predicted polypeptide is 268 amino acids long and has a size of 28.6 kDa (Fig. 4A). The sequence homology search using the GenBank and EMBL database indicated that the encoded protein had a significant homology to C/EBP in the carboxy-terminal part containing the basic region and the leucine zipper motif (Fig. 4B), and thus we named this factor CELF (C/EBP-like factor). In the carboxy-terminal 76-amino acid region, CELF had 58% identity with C/EBP, 62% identity with IL-6DBP, and 39% identity with IgE/BP-1. However, upstream of this region CELF had no apparent homologies to these proteins, although the amino-terminal part was relatively rich in prolines (16%; 30 out of 192 amino acids) like other C/EBP gene family proteins. CELF did not have any significant homology to CRE-binding proteins such as CREB and ATF2 except for the leucine zipper motif. Unlike CREB and ATF2, CELF did not have a typical phosphorylation site by protein kinase A. Because CELF is a novel member of the C/EBP gene family, we decided to further characterize this protein.

**CELF and ATF2 Bind to the CRE Region of the PPT-A Gene**—To analyze the DNA binding specificity of CELF, we cloned the cDNA in pGEMEX1 (Promega), a T7 expression vector, and expressed the fusion protein in *Escherichia coli*. The fused protein contained the amino-terminal 26-kDa fragment of the T7 gene 10 product and the 28.6-kDa CELF fragment, and thus, the size of this fused protein should be approximately 55 kDa. We obtained the product of this size, and after further purification from the SDS-polyacrylamide gel (Fig. 5A), we subjected it to the DNA binding experiment. We also made the ATF2 fusion protein and the control gene 10 protein for the DNA binding analysis. As shown in Fig. 5B, the CELF fusion protein protected
The deduced primary structure of the rat counterpart of ATP2. The cDNA insert of Al was sequenced. The leucine residues of the leucine zipper motif are boxed.

Transcription Factor CELF

Fig. 3. Nucleotide sequence and the deduced primary structure of rat counterpart of ATP2. The cDNA insert of Al was sequenced. The leucine residues of the leucine zipper motif are boxed.
the CRE region of the PPT-A gene (lane 2). This protection pattern was identical with that of the ATP2 fusion protein (lane 3), indicating that CELF and ATF2 recognize the same target. The mixture of these two proteins did not appear to alter the protection region (lane 4). The control gene 10 protein did not protect any specific regions (lane 2).

**Transcription Activation by CELF**—To characterize the functional effects of CELF, we performed cell-free transcription experiments. The activation effect of CELF was tested using the nuclear extracts prepared from A431 human epidermoid carcinoma cell lines (Kageyama et al., 1988). As shown in Fig. 6, the CELF fusion protein exhibited a 6-fold stimulation of the PPT-A gene expression (compare lanes 1 and 2), while this factor showed almost no activation on the enhancerless SV40 early promoter (compare lanes 3 and 4), which does not contain CELF-binding sites. However, insertion of the CRE of the PPT-A gene into the enhancerless SV40 early promoter resulted in 2-fold induction of transcription (compare lanes 5 and 6). These results indicate that CELF is involved at the transcriptional level and acts as an activator by interacting with the CRE region. The active CELF fusion protein was synthesized in E. coli, where no effects of protein kinase A are known in contrast to the eukaryotic mechanism (for review see de Crombrugghe et al. (1984)). Therefore, the cell-free transcription data shown above further suggest that CELF could function as a constitutive activator irrespective of protein kinase A phosphorylation.

To confirm this point, we next analyzed the phosphorylation effects by protein kinase A. We treated the CELF fusion protein with the catalytic subunit of protein kinase A in the presence of [γ-32P]ATP. However, we could not detect any labeling by protein kinase A (data not shown). CELF thus does not seem to serve as a substrate for protein kinase A. Consistent with this finding, the kinase-treated CELF did not show any significant change compared with the untreated CELF (compare lanes 7 and 8). These results demonstrate that CELF supports the constitutive activation rather than cAMP-inducible transcription.

**Tissue Distribution of CELF**—The tissue distribution of CELF mRNA was analyzed by a Northern blot experiment (Fig. 7). We detected two main bands with molecular sizes of 8 and 1.3 kb. We found the two mRNA species of CELF in all tissues we tested, indicating that this gene is ubiquitously expressed. We also found that CELF mRNA was more abundant in testes and adrenal than in other tissues. These results suggest that CELF is widely utilized by various genes and could contribute to complex transcriptional regulation through CRE sequences.

**DISCUSSION**

**CELF Constitutively Activates Transcription**—cAMP response elements are present in various genes and are involved
in cAMP-induced gene expression. To date several factors that recognize a CRE have been identified. One such example is CREB, which activates transcription after phosphorylation by protein kinase A (Gonzalez and Montminy, 1989; Ribasovic et al., 1988). In this study, we have isolated a cDNA encoding a new transcription factor that binds to the CRE region. This factor does not seem to require protein kinase A-mediated phosphorylation for this activity and constitutively activates proteins. In this situation, CELF could function as a modulator of the CRE.

The CRE supports at least two different transcription levels: the basal level transcription via the CRE. However, it is possible that CELF could support regulatory pathways of other effects such as protein kinase C.

Because both CELF and other CRE-binding proteins are ubiquitous factors and share similar structure characteristics, the balance of these proteins should play a critical role in regulating various genes containing CRE sequences. As we have noted, CELF has a leucine zipper structure, a motif required for dimer formation (Landschulz et al., 1987). We thus speculate that CELF is involved in basal level transcription via the CRE. However, it is possible that CELF could support regulatory pathways of other effects such as protein kinase C.

Another intriguing question concerns the significance of the CRE for PPT-A gene regulation. The expression of substance P is controlled by various factors such as NGF (Lindsay and Harmar, 1989). Accumulating evidence implicates that the NGF action involves at least two signal transduction systems, cAMP and Ca2+ / phosphatidylinositol pathways (Cremins et al., 1986). Thus, NGF stimulation of the substance P synthesis may occur via the cAMP-induced transcription of the PPT-A gene. Moreover, recent studies suggest that a calcium response element that induces transcription by membrane depolarization is indistinguishable from a CRE (Sheng et al., 1990). Therefore, NGF-induced Ca2+ mobilization may also contribute to PPT-A gene activation through the CRE. However, Ca2+ regulation of the PPT-A gene needs more careful analysis, because it has been reported that membrane depolarization decreases the PPT-A mRNA level in ganglion cells (Black et al., 1987).

Our report has now established the basis for understanding the transcriptional mechanisms of the PPT-A gene, and these studies should also help analyze regulatory mechanisms of other neuropeptide and transmitter expressions.

**FIG. 7. Northern blot analysis of CELF.** 10 μg each of poly (A)+ RNA of the whole brain (lane 1), liver (lane 2), kidney (lane 3), adrenal (lane 4), and testis (lane 5) were analyzed with a 2.1-kb EcoRI fragment of X2 as a probe. Two RNA species of CELF are shown by arrows. The positions of 28 and 18 S ribosomal RNAs are indicated on the left.
Acknowledgments—We are grateful to Drs. Alfred Johnson and Ira Pastan for providing invaluable A431 nuclear extracts and to Dr. Marc Montminy for the generous gift of the expression vector of the catalytic subunit of protein kinase A.

REFERENCES

Friedman, A. D., and McKnight, S. L. (1990) Genes & Dev. 4, 1416–1426