

# DNA Binding Specificity of the CCAAT/Enhancer-binding Protein Transcription Factor Family\*

(Received for publication, September 20, 1995, and in revised form, November 16, 1995)

Shigehiro Osada, Hiroko Yamamoto, Tsutomu Nishihara, and Masayoshi Imagawa†

From the Department of Environmental Biochemistry, Faculty of Pharmaceutical Sciences, Osaka University,  
1-6 Yamada-Oka, Suita, Osaka 565, Japan

CCAAT/enhancer-binding protein (C/EBP) transcription factor family members are related by a high degree of amino acid sequence identity to the basic leucine zipper DNA-binding domain and show distinct but overlapping patterns of tissue- and stage-restricted expression. Although C/EBP $\alpha$  and C/EBP $\beta$  have been shown to recognize a consensus sequence derived from regulatory elements in virus and acute-phase response genes, the potential for more subtle differences in the binding preference of the C/EBP family has not been previously addressed. The consensus sequence of C/EBP $\delta$  has not been reported. By using the method of polymerase chain reaction-mediated random site selection to assess the DNA binding specificity of the C/EBP family in an unbiased manner, we demonstrated the sequence preferences for C/EBP family members. With small variations, these C/EBP family members showed similar sequence preferences, and the consensus sequence was identified as RTTGCGYAAAY (R = A or G, and Y = C or T). The phosphorylation of C/EBP $\delta$  by casein kinase II increased the binding activity, but did not affect the binding specificity, whereas it was reported that the phosphorylation of C/EBP $\alpha$  and C/EBP $\beta$  decreased the binding affinity. The specificity of action of C/EBP family members may be derived from the characteristics of each factor, including the expression profiles, the DNA binding affinities, the cofactors, and so on, in addition to the DNA binding specificities.

Many transcription factors have been found to be members of highly related multifactor families, and thus, their specificity of action must be addressed in order to ascertain their respective functions. CCAAT/enhancer-binding protein (C/EBP)<sup>1</sup> family members are among the basic leucine zipper transcription factors, and they bind to specific DNA sequences as dimers. Six C/EBP proteins (designated C/EBP $\alpha$ , C/EBP $\beta$ , C/EBP $\gamma$ , C/EBP $\delta$ , C/EBP $\epsilon$ , and CHOP 10) have been identified, and C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$  have been studied in detail (1–12).

C/EBP $\alpha$ , previously termed C/EBP, was identified originally as a heat-stable protein present in soluble extracts of rat liver nuclei and having sequence-specific DNA binding activity. Pu-

rified C/EBP $\alpha$  selectively recognized CCAAT homologies and enhancer core sequences, implying that it might be a transcriptional regulatory protein. Alignment of the avian retroviral binding sequences with the published binding sites for C/EBP $\alpha$  in two CCAAT boxes and in the simian virus 40, polyoma, and murine sarcoma virus enhancers suggested TKNNGYAAK (K = T or G, and Y = C or T) as a consensus sequence for binding of C/EBP $\alpha$  (13). Several studies showed that C/EBP $\alpha$  trans-activates liver- and adipose-specific genes as well as virus genes, while C/EBP $\beta$  and C/EBP $\delta$  were cloned as nuclear factors that regulate acute-phase response genes and liver-specific genes (2–4, 7). Akira *et al.* (2) reported that competition analysis of the binding of NF-IL6 (human C/EBP $\beta$ ) with the published sequences to which C/EBP-like proteins bound revealed that NF-IL6 and C/EBP $\alpha$  recognized the same nucleotide sequences and that the best fit was the consensus TKNNGNAAK for NF-IL6. The consensus sequence of C/EBP $\delta$  has not been reported. The basic leucine zipper regions of C/EBP isoforms show high similarity, and some *cis*-elements are recognized by each of the C/EBP isoforms (4, 6–8). Each isoform, however, shows distinct but overlapping patterns of tissue- or stage-restricted expression (2–4, 6–10). Therefore, it is important to determine the specificity of the C/EBP family. The polymerase chain reaction (PCR)-mediated random site selection method has been adopted to distinguish the binding sites of the transcription factor family members (14–19).

Modification of a transcription factor, such as by phosphorylation, glycosylation, and reduction-oxidation, affects its binding activity and function. Phosphorylation of C/EBP $\alpha$  by protein kinase C results in an attenuation of binding, and modification of C/EBP $\beta$  by protein kinases A and C results in an inhibition of DNA binding (20, 21). In the case of C/EBP $\delta$ , dephosphorylation severely decreases the DNA binding ability *in vitro*, although neither kinase nor phosphatase for this protein has been identified yet *in vivo* (22).

Utilizing a binding site selection procedure, we show here that the bacterially expressed C/EBP proteins have virtually identical binding specificity and that the phosphorylation of C/EBP $\delta$  by casein kinase II increases the binding activity, but does not affect the binding specificity.

## MATERIALS AND METHODS

**Expression and Purification of Bacterially Expressed Proteins—**C/EBP $\alpha$  was expressed in *Escherichia coli* by the glutathione *S*-transferase fusion protein system (Pharmacia Biotech Inc.). A *Sma*I-*Pst*I fragment derived from pMSV-C/EBP $\alpha$  (23) and encoding the DNA-binding domain was subcloned into a pGEX-3X expression vector. The recombinant plasmid was transformed into M15(pREP4) (QIAGEN Inc.). The transformant was grown overnight at 30 °C in LB medium containing 100  $\mu$ g/ml ampicillin and 25  $\mu$ g/ml kanamycin. The culture was then diluted 100-fold and grown to A<sub>600</sub> = 0.4; at that time, isopropyl- $\beta$ -D-thiogalactopyranoside was added to a final concentration of 0.3 mM. Then, the cells were allowed to grow for an additional 4 h, harvested by centrifugation, suspended in 10% of the original culture volume of 0.1 M TM buffer (50 mM Tris (pH 7.9), 1 mM EDTA, 1 mM dithiothreitol, 12.5

\* This work was supported in part by grants from the Ministry of Education, Science, Sports, and Culture, Japan and from the Asahi Glass Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed. Tel.: 81-6-879-8241; Fax: 81-6-879-8244; E-mail: imagawa@phs.osaka-u.ac.jp.

<sup>1</sup> The abbreviations used are: C/EBP, CCAAT/enhancer-binding protein; PCR, polymerase chain reaction; AlbD, albumin D site; GPS1, glutathione transferase P silencer 1; CRP, C-reactive protein.

mM MgCl<sub>2</sub>, 20% glycerol, 0.1 M KCl), and disrupted by sonication. Following centrifugation at 13,500 rpm for 15 min, the supernatant was loaded onto a glutathione-Sepharose column equilibrated with 0.1 M TM buffer, washed with the same buffer, and eluted with 10 mM glutathione in 0.1 M TM buffer.

The nucleotide sequence of C/EBP $\beta$  is identical to that of Silencer Factor B previously reported (5). For production of the C/EBP $\beta$  DNA-binding domain, we used Y1089 lysogen, which contains the DNA-binding domain of C/EBP $\beta$ /Silencer Factor B cDNA, but lacks 84 amino-terminal amino acids. Bacterial C/EBP $\beta$  was produced and purified as described previously (5).

C/EBP $\delta$  was expressed by the histidine fusion protein system (QIAGEN Inc.). The *XhoI*-*DraI* fragment derived from the rat genomic C/EBP $\delta$  gene<sup>2</sup> and encoding the DNA-binding domain was subcloned into a pQE-30 expression vector. The recombinant plasmid was transformed into M15(pREP4). The transformant was grown under the same conditions as for C/EBP $\alpha$  with a slight modification. In brief, the protein was induced by isopropyl- $\beta$ -D-thiogalactopyranoside when A<sub>600</sub> reached 0.8. The cells were suspended in 10% of the original culture volume of buffer A (6 M guanidine hydrochloride, 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, 0.01 M Tris (pH 8.0)) and stirred for 1 h at room temperature. The supernatant obtained by centrifugation at 13,500 rpm for 15 min was loaded onto a Ni<sup>2+</sup>-nitrilotriacetic acid column equilibrated with buffer A, washed with the same buffer, and eluted with buffer E (8 M urea, 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, 0.01 M Tris (pH 4.5)). The eluate was dialyzed against 0.1 M HM buffer (25 mM HEPES (pH 7.8), 1 mM dithiothreitol, 12.5 mM MgCl<sub>2</sub>, 20% glycerol, 0.1 M KCl), and the resultant supernatant was used for DNA binding analysis.

**Random Binding Site Selection**—The following synthetic oligonucleotides (Primers 1–4) were used for the binding site selection (*Bam*HI and *Xho*I restriction sites are underlined).

```

5'-AGAGCCACTTCTCTCAACGGATCC-3'   Primer 1
          5'-CTCAACGGATCCGTC-3'   Primer 3
5'-AGAGCCACTTCTCTCAACGGATCCGTC----- (N)16-----
          -----GTCCTCGAGACTGTGAGTCGTCTGAC-3'
Primer 4      3'-CAGGAGCTCTGACAG-5'
Primer 2      3'-GAGCTCTGACAGTCAGCAGACTG-5'
PRIMERS 1-4

```

Double-stranded molecules were generated by annealing oligonucleotides containing 16 random nucleotides with Primers 1 and 2 and then extending by *Taq* DNA polymerase (Promega). Enrichment for binding sites was performed by the filter binding method (19, 24). The binding mixture (35  $\mu$ l) contained purified bacterially expressed protein (1  $\mu$ g of C/EBP $\alpha$ , 7  $\mu$ g of C/EBP $\beta$ , or 0.6  $\mu$ g of C/EBP $\delta$ ), 0.3  $\mu$ g of double-stranded random oligonucleotide, 1.4  $\mu$ g of poly(dI-dC), and 3.5  $\mu$ l of 10  $\times$  binding buffer (100 mM Tris-HCl (pH 7.5), 50% glycerol, 10 mM dithiothreitol, 10 mM EDTA). Each mixture was incubated for 30 min at room temperature. Thereafter, this solution was passed slowly through a presoaked nitrocellulose filter (Schleicher & Schuell, BA85). The filter was washed three times with 3 ml of 1  $\times$  binding buffer before the bound oligonucleotides were eluted with 100  $\mu$ l of elution buffer containing 20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 20 mM NaCl, and 0.1% SDS. The eluate was phenolized and then amplified by PCR using Primers 1 and 2. The amplified products (0.3  $\mu$ g) were put back into a binding reaction, and the procedure was repeated. After five rounds (in the case of C/EBP $\alpha$  and C/EBP $\beta$ ) or seven rounds (in the case of C/EBP $\delta$ ) of the enrichment, the amplified products were end-labeled with [ $\gamma$ -<sup>32</sup>P]ATP by using T4 polynucleotide kinase and used as a probe. The labeled amplified products were incubated with bacterially expressed C/EBP proteins, and protein-DNA complexes were separated on 6% (in the case of C/EBP $\alpha$  and C/EBP $\beta$ ) or 4% (in the case of C/EBP $\delta$ ) native polyacrylamide gels. Bound DNA was eluted with a solution of 0.5 M ammonium acetate, 1 mM EDTA, 0.1% SDS, 10% methanol, and 50  $\mu$ g/ml proteinase K. The eluate was phenolized and then amplified by PCR using Primers 3 and 4. Amplified products were digested with *Bam*HI and *Xho*I and subcloned into pBluescript KS (Stratagene). Clones were sequenced by the dideoxy chain termination method (25).

**Gel Shift Analysis**—Sequences of the oligonucleotides for gel shift analysis was as follows (only upper strands are shown): pal, 5'-CTAGCATATTGCGCAATATGC-3'; m+/-1, 5'-CTAGGCATATTGGCCA-

ATATGC-3'; m+/-2, 5'-CTAGGCATATTCGGAATATGC-3'; m+/-3, 5'-CTAGGCATATAGCGCTATATGC-3'; m+/-4, 5'-CTAGGCATAATGCGCATATATGC-3'; m+/-5, 5'-CTAGGCATTTTGGCGAAATATGC-3'; +3-G, 5'-CTAGGCATATTGCGCGATATGC-3'; +3-T, 5'-CTAGGCATATTGCGCTATATGC-3'; +3-C, 5'-CTAGGCATATTGCGCCATATGC-3'; AlbD, 5'-CTAGTGGTATGATTTTGTAAATGGGGTA-3'; GPS1, 5'-CTAGAGAGGTTGGTAAATAGGGATGG-3'; and CRP, 5'-CTAGCA-TAGTGGCGCAAACCTCCCTTA-3'. The protein fraction (6.25  $\mu$ l) containing the purified protein or the phosphorylation reaction mixture (see below) was mixed with the same volume of 20 mM Tris-HCl (pH 7.5), 10% glycerol, 2 mM dithiothreitol, 10 mM EDTA, 0.5  $\mu$ g of poly(dI-dC), 1  $\mu$ g of bovine serum albumin, 0.1% Nonidet P-40 (in the case of the phosphorylation reaction mixture), and 0.8 ng of labeled probe. The binding reaction was performed at room temperature for 30 min. Each reaction mixture was loaded on a nondenaturing polyacrylamide gel, electrophoresed at 150 V for 1 h, fixed with 10% methanol and 10% acetic acid, and autoradiographed.

**Phosphorylation and Dephosphorylation of C/EBP $\delta$** —For phosphorylation of C/EBP $\delta$ , purified bacterially expressed C/EBP $\delta$  was incubated with 0.2 milliunits of casein kinase II (Boehringer Mannheim) for 1 h at 37 °C in 25  $\mu$ l of the phosphorylation reaction mixture containing 2.5  $\mu$ g of bovine serum albumin, 10 mM dithiothreitol, 40 mM HEPES (pH 8.4), 3 mM MgCl<sub>2</sub>, and 4 mM ATP. For the dephosphorylation experiments, 0.6 units of bacterial alkaline phosphatase (TOYOBO) was added directly to the phosphorylation reaction mixture. These samples were mixed with 63 mM Tris (pH 6.8), 10% glycerol, 2% SDS, 0.72 M  $\beta$ -mercaptoethanol, and 0.02% bromophenol blue; boiled; and then electrophoresed on 15% SDS-polyacrylamide gels.

## RESULTS

**Selection of DNA-binding Sites for C/EBP Family Members**—We selected oligonucleotides that could be recognized by C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$  from the double-stranded N<sub>16</sub> randomized oligonucleotides by using bacterially expressed proteins. The optimal C/EBP $\alpha$ -binding site consists of directly abutted, dyad half-sites bearing the pentanucleotide sequence 5'-ATTGC-3' (26). Three-hundred nanograms of randomized oligonucleotide contains  $\sim 4 \times 10^{12}$  molecules, sufficient to ensure that all possible combinations of the 10-nucleotide randomized region ( $4^{10} = 1 \times 10^6$ ) would be represented in each binding reaction. The PCR-amplified DNA recovered after each round of selection was subjected to gel shift analysis, and recovery of DNA increased with each selection (data not shown). DNAs recovered after five rounds of selection (in the case of C/EBP $\alpha$  and C/EBP $\beta$ ) or seven rounds (in the case of C/EBP $\delta$ ) were cloned into a plasmid vector and sequenced. The sequences of 81, 99, and 83 DNAs were independently obtained for C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$ , respectively (data not shown). Sequences of cloned oligonucleotides were optimally aligned by inspection, a task greatly facilitated by the presence within most sequences of the pentanucleotide sequence 5'-ATTGC-3'. This motif, used to derive a preliminary consensus sequence, was found in 59.3% (C/EBP $\alpha$ ), 58.6% (C/EBP $\beta$ ), and 32.5% (C/EBP $\delta$ ) of the oligonucleotides. This preliminary consensus sequence was used as a guide to align optimally the remaining sequences that contained 5'-ATTGC-3'-like sequence. A condensed list of sites selected by each protein is presented in a tabular form in Fig. 1A. The alignment of the binding site selection data is summarized according to frequency of the nucleotide at each position of a composite pentanucleotide as well as flanking nucleotides, and the consensus binding site for each member of the C/EBP family is established (Fig. 1B). The consensus binding site, 5'-RTTGC-GYAA-3' (R = A or G, and Y = C or T), is similar in all three proteins, with small differences. The consensus binding site revealed that C/EBP isoforms could bind to palindrome or palindrome-like sequences composed of two 5-base pair half-sites. The bases in the sequence recognized by C/EBP family members are numbered 1–5. The positive number refers to the position in the right half-site, and the negative to that in the left half-site.

<sup>2</sup> T. Yamada, T. Nishihara, and M. Imagawa, unpublished data.

**A**

	C/EBP $\alpha$ (%)	C/EBP $\beta$ (%)	C/EBP $\delta$ (%)
ATTGC	59.3	58.6	32.5
GTTGC	17.3	21.2	13.3
CTTGC	7.4	7.1	14.5
TTTGC	0	2.0	4.9
ATTGT	1.2	2.0	1.2
GTTGT	1.2	0	3.6
ATTGG	0	1.0	2.4
ATTGA	0	1.0	0
ATTAC	8.6	5.1	13.3
GTTAC	3.7	2.0	7.2
ATTAT	0	0	1.2
GTTAG	1.2	0	0
GTTTC	0	0	2.4
ATGTC	0	0	2.4
GTGTC	0	0	1.2

**B**

C/EBP $\alpha$												
G (%)	29.6	23.5	0	0	86.4	1.2	80.2	0	0	0	16.0	23.5
A (%)	17.3	69.1	0	0	13.6	0	14.8	17.3	90.1	100	0	29.6
T (%)	34.6	0	100	100	0	2.5	1.2	51.9	0	0	51.9	24.7
C (%)	18.5	7.4	0	0	0	96.3	3.7	30.9	9.9	0	32.1	22.2
consensus	A/G T T G C G T/C A A T/C											
C/EBP $\beta$												
G (%)	34.3	23.2	0	0	92.9	1.0	77.8	1.0	0	0	9.1	33.3
A (%)	13.1	67.7	0	0	7.1	1.0	21.2	17.2	91.9	100	1.0	22.2
T (%)	26.3	2.0	100	100	0	2.0	1.0	35.4	0	0	63.6	21.2
C (%)	26.3	7.1	0	0	0	96.0	0	46.5	8.1	0	26.3	23.2
consensus	A/G T T G C G C/T A A T/C											
C/EBP $\delta$												
G (%)	21.7	28.9	0	3.6	75.9	4.8	63.9	0	0	0	18.1	26.5
A (%)	22.9	50.6	0	0	18.1	1.2	25.3	28.9	66.3	98.8	4.8	24.1
T (%)	27.7	4.8	100	96.4	6.0	6.0	3.6	48.2	2.4	0	42.2	30.1
C (%)	27.7	15.7	0	0	0	88.0	7.2	22.9	31.3	1.2	34.9	31.3
consensus	A/G T T G C G T/AC A A T/C											
position	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5		

FIG. 1. Compilation of sequences selected by C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$ . Bacterially expressed proteins were used for selection of random oligonucleotides. A, list of pentamer sequences selected by C/EBP isoforms. The percentages of pentamer sequences selected by each protein are shown. B, alignment of the binding sequences selected. The frequency of nucleotide selection at each position of a composite pentanucleotide as well as flanking nucleotides is indicated as percent. The consensus binding sequences for each isoform are also shown. The bases in the consensus sequence are numbered consecutively from 1 to 5 (the values are positive for the right half-site and negative for the left-half site).

**DNA Binding Specificity of C/EBP Proteins**—To determine whether C/EBP family members recognize palindrome sequence or each base in a consensus binding site, gel shift analysis using the palindrome-type double point mutants as competitors was performed. As a probe, we adopted the palindrome sequence that contains 5'-ATTGCGCAAT-3'. DNA-protein complexes with or without a competitor are shown in Fig. 2A. C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$  showed the same binding specificity. As the mutant oligonucleotides (m+/-3 and m+/-4) completely failed to compete the binding, the bases at positions +/-3 and +/-4 were deduced to be the most important binding sequences for C/EBP isoforms. After position +/-3 and +/-4 bases, position +/-2 bases are more significant than position +/-1 and +/-5 bases.

The selection data revealed that sequences previously identified as binding sites for some of the proteins were rarely selected. We examined the hypothesis that C/EBP isoforms could bind to mutant oligonucleotides at position +3, one of the most important bases for binding. A palindrome sequence including 5'-ATTGCGCAAT-3' was used as a probe, and the mutant oligonucleotides were adopted as competitors (Fig. 2B). The palindrome sequence completely competed binding to the probe at low concentration. The mutant oligonucleotides also inhibited binding to the specific probe at higher concentration.

The mutant +3-C showed more effective competition than the other two competitors, +3-G and +3-T. These data revealed that C/EBP isoforms could bind to the mutant oligonucleotides at position +3 (+3-G, +3-T, and +3-C).

**Phosphorylation of C/EBP $\delta$  by Casein Kinase II**—C/EBP isoforms show distinct but overlapping patterns of restricted expression (2-4, 6-10), but they have the same DNA binding specificities as revealed in Figs. 1 and 2. Phosphorylation and dephosphorylation of transcription factors have been shown to be crucial to the ability of these proteins to bind to their cognate DNA sequences. C/EBP $\alpha$  and C/EBP $\beta$  are phosphorylated by protein kinase C and/or protein kinase A, which results in an attenuation of binding *in vitro* (20, 21). However, dephosphorylation of the liver nuclear extracts from the acute phase-induced rabbit reduces the DNA binding ability of endogenous C/EBP $\delta$  (22). Manak *et al.* (27) reported that casein kinase II enhances the DNA binding activity of the serum response factor. We investigated whether C/EBP $\delta$  could be a substrate in phosphorylation by casein kinase II *in vitro*. Purified bacterially expressed C/EBP $\delta$  was incubated with casein kinase II and separated by SDS-polyacrylamide gel electrophoresis. The migration of untreated C/EBP $\delta$  was faster than that of C/EBP $\delta$  incubated with casein kinase II, whereas it was the same as that of C/EBP $\delta$  incubated with casein kinase II and bacterial

alkaline phosphatase (Fig. 3A). Furthermore, the labeled palindrome probe containing 5'-ATTGCGCAAT-3' was incubated with several amounts of nonphosphorylated or phosphorylated C/EBP $\delta$  and then separated by native polyacrylamide gel electrophoresis. The binding activity of C/EBP $\delta$  was enhanced about three times by phosphorylation (Fig. 3B).

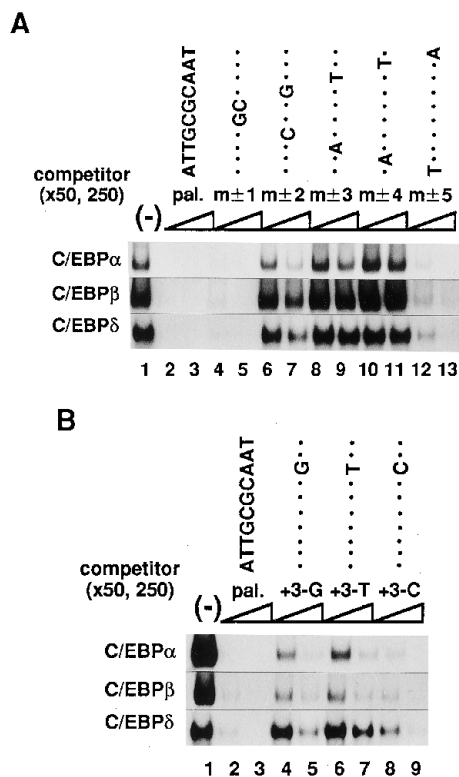
**DNA Binding Specificity of Phosphorylated C/EBP $\delta$** —To compare the DNA binding specificities of phosphorylated and nonphosphorylated C/EBP $\delta$ , a gel shift assay using double point and single point (+3) mutant oligonucleotides was performed. The DNA binding specificity of phosphorylated C/EBP $\delta$

resembled that of nonphosphorylated C/EBP $\delta$  (Figs. 2 and 4). The bases at positions  $\pm 3$  and  $\pm 4$  were the most important for phosphorylated C/EBP $\delta$  to recognize its binding site, followed by position  $\pm 2$  bases and then position  $\pm 1$  and  $\pm 5$  bases (Fig. 4A). As shown in Fig. 4B, the experiment using the single point (+3) mutant oligonucleotides indicated that phosphorylated C/EBP $\delta$  also could bind to oligonucleotides mutated at position +3 (+3-G, +3-T, and +3-C), which is among the most important positions for C/EBP binding.

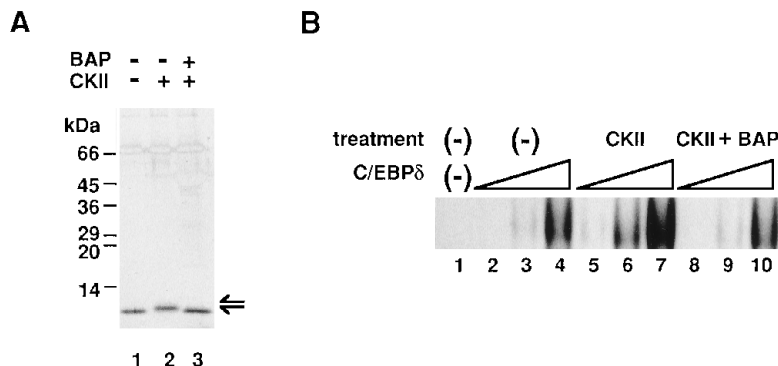
**Relative Affinities of Bacterially Expressed C/EBP Isoforms and Phosphorylated C/EBP $\delta$  for Native C/EBP Isoform-binding Sites**—Although several binding sites of C/EBP isoforms have been reported, many of them have a few bases different from the consensus binding site. For example, those for AlbD and GPS1 contain TT and AA at positions  $\pm 3$  and  $\pm 4$ , but differ from the consensus binding sites at 1 or 2 base(s) between positions  $-2$  and  $+2$  (28, 29). On the other hand, the acute-phase response element in the CRP gene is identical to the consensus sequence between positions  $-2$  and  $+2$ , but position +3 is C (30). We observed the affinities of C/EBP isoforms for several native binding sites (Fig. 5). AlbD showed a higher affinity than GPS1 and CRP, but a lower affinity than the consensus binding site (Fig. 2, A and B, lanes 2 and 3; and Fig. 5). The same affinity was observed between GPS1 and CRP. Such relationships were also observed in all isoforms of nonphosphorylated C/EBP and phosphorylated C/EBP $\delta$ .

#### DISCUSSION

A gel shift selection and PCR amplification using randomized oligonucleotides and bacterially expressed proteins allowed us to identify the preferred DNA-binding sites for C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$ . Most of the selected clones, 76.6% (C/EBP $\alpha$ ), 79.8% (C/EBP $\beta$ ), and 45.8% (C/EBP $\delta$ ), contained 5'-RTTGC-3' ( $R = A$  or  $G$ ) sequence (Fig. 1A). The consensus binding site for each member of the C/EBP family was determined by an alignment of the binding site selection data, which were summarized by the frequency of nucleotide selection at each position of a composite pentanucleotide as well as flanking nucleotides. The consensus binding site, 5'-RTTGCYAAAY-3' ( $R = A$  or  $G$ , and  $Y = C$  or  $T$ ), for all three proteins is similar, except for small differences. Many sequences of the binding sites selected by C/EBP $\delta$  are not identical to the consensus binding sites. The hierarchy of DNA binding affinities for the consensus sequence of C/EBP isoforms is C/EBP $\beta$  > C/EBP $\alpha$  > C/EBP $\delta$  (data not shown) (7). In the case of C/EBP $\alpha$  and C/EBP $\beta$ , there is a difference between the high and low affinity binding sites, with the former being mainly selected. However, in the case of C/EBP $\delta$ , because of



**FIG. 2. DNA binding specificity of C/EBP family members.** Double-stranded oligonucleotides containing palindrome consensus sequence for C/EBP binding were incubated with bacterially expressed C/EBP proteins in the absence (lane 1) or presence (A, lanes 2–13; B, lanes 2–9) of increasing amounts of the competitor for gel shift analyses. Each competitor was present in either 50-fold excess (even-numbered lanes) or 250-fold excess (odd-numbered lanes) relative to the probe. Competitor sequences (A, double point mutants; B, position +3 mutants) are shown on top of the lanes.



**FIG. 3. Effect of phosphorylation on the binding ability of C/EBP $\delta$  by casein kinase II.** A, bacterially expressed C/EBP $\delta$  was incubated without casein kinase II (CKII) (lane 1), with casein kinase II (lane 2), and with casein kinase II and bacterial alkaline phosphatase (BAP) (lane 3) and then analyzed by SDS-polyacrylamide gel electrophoresis. Proteins were detected by silver staining. The arrows designate phosphorylated (upper) and nonphosphorylated (lower) C/EBP $\delta$ . B, shown is the gel shift analysis of phosphorylated C/EBP $\delta$ . Increasing amounts (0.11, 0.33, and 1 ng) of C/EBP $\delta$  (nontreatment (lanes 2–4), treatment with casein kinase II (lanes 5–7), and treatment with casein kinase II and bacterial alkaline phosphatase (lanes 8–10)) were incubated with the probe (pal) and separated by native polyacrylamide gel electrophoresis.

small differences in affinity, many of the low affinity binding sites were also selected. It was reported that C/EBP $\alpha$  and C/EBP $\beta$  could recognize a consensus sequence derived from regulatory elements in virus and acute-phase response genes, TKNNGYAAK or TKNNGNAAK ( $K = T$  or  $G$ , and  $Y = C$  or  $T$ ), respectively (2, 13). These consensus sequences are not palindromic, although they resemble our results. As our system included all possible combinations of the 10-nucleotide randomized sequence, the selected binding sites were unbiased. Our data revealed that C/EBP isoforms could recognize the palindrome sequence. The precise nucleotide identities of positions  $\pm 3$  and  $\pm 4$  were the most critical determinants of recognition specificity and affinity. Position  $\pm 2$  bases were the third most important. Bases at positions  $\pm 1$  and  $\pm 5$  were also selected weakly, indicating that C/EBP isoforms interact at these positions with a moderate degree of sequence preference. The crystal structure of the GCN4 and AP-1 sites revealed that one of the bases near the center of the AP-1 site is not contacted by the basic leucine zipper structure (31). Because position  $\pm 1$  bases are located in the center, these bases are not in contact with the basic leucine zipper structure and show higher flexibility compared with positions  $\pm 2$ ,  $\pm 3$ , and  $\pm 4$ . Many binding sites of C/EBP family members have been previously identified. Most of them, however, are not completely coincident with the consensus sequence reported here. When significant base(s) for the DNA binding are not identical to the consensus sequence and other bases are conserved, C/EBP members could weakly bind to the DNA. Binding site sequence that is more similar to the consensus se-

quence, especially at positions  $\pm 3$  and  $\pm 4$ , showed a much higher affinity. We used bacterially expressed DNA-binding domain fusion proteins, such as glutathione *S*-transferase,  $\beta$ -galactosidase, and histidine tag. Gel shift analysis using mutant binding sites revealed that the same binding specificities were shown regardless of the kind of fusion protein (data not shown). Sequence-specific transcription factors are composed of structural domains for DNA binding and transcriptional regulatory functions, and DNA binding specificities are usually determined by the DNA-binding domain. However, it seems possible that the transcriptional regulatory domain affects the DNA binding specificities. These possibilities remain to be resolved.

The binding activity of C/EBP isoforms was modified by phosphorylation. It is known that phosphorylation of the DNA-binding domain of C/EBP $\alpha$ , containing Ser<sup>299</sup> in the basic region of the basic leucine zipper structure, which is the phosphorylation site by protein kinase C, results in attenuation of DNA binding (20). Phosphorylation of C/EBP $\beta$  at Ser<sup>240</sup> within the DNA-binding domain by protein kinase A results in a decrease in its binding to DNA, while phosphorylation of Ser<sup>105</sup> within the transcriptional activation domain has no effect on DNA binding (21). In the case of C/EBP $\delta$ , dephosphorylation of the nuclear extracts reduces the DNA binding ability of endogenous C/EBP $\delta$  *in vitro*, although the phosphorylation site and the kinase and/or phosphatase for this protein have not been identified yet *in vivo* (22). When C/EBP $\delta$  was phosphorylated by incubation with casein kinase II, DNA binding activity was promoted (Fig. 3). Casein kinase II phosphorylates serine residues located in acidic regions of proteins (32). Five serine residues exist in the C/EBP $\delta$  DNA-binding domain, and one of them, Ser<sup>227</sup>, exists in acidic regions (Glu-Leu-Ser<sup>227</sup>-Ala-Glu-Asn-Glu), although it is unclear whether Ser<sup>227</sup> is actually phosphorylated by casein kinase II *in vitro* and *in vivo*. We could not rule out the possibility that other sites of phosphorylation on C/EBP $\delta$  may also affect the structure and therefore the specificity and affinity of the protein for its DNA targets *in vivo*.

We found that the consensus binding sites for the C/EBP family members are similar to each other, suggesting that same target genes might be regulated in the cells. The function of the C/EBP-binding site is classified in three ways. AlbD and CRP act as a constitutive and an inducible enhancer, respectively (3, 4). On the other hand, GPS1 functions as a silencer (29). The order of the affinity of the three binding sites depends on the similarity to the consensus sequences. These sites show different functions, but they behave similarly when binding. Considering the gene expression regulated by the C/EBP family, the amount of expression and the affinity for DNA of C/EBP proteins are important factors. In the cell, many parameters must collaborate to define the function of a given C/EBP-bind-

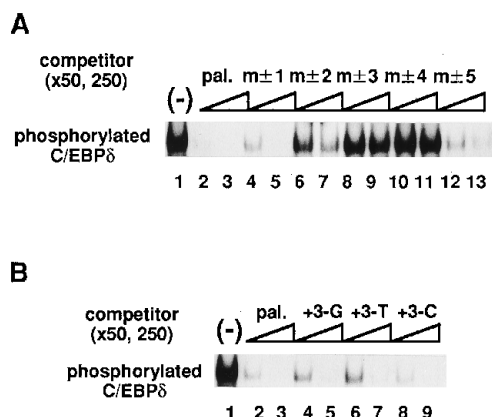


FIG. 4. DNA binding specificity of phosphorylated C/EBP $\delta$ . Probes containing palindrome consensus sequence were incubated with C/EBP $\delta$  phosphorylated by casein kinase II in the absence (lane 1) or presence (A, lanes 2–13; B, lanes 2–9) of increasing amounts of the competitor (50- and 250-fold molar excesses) for gel shift analysis. Competitor sequences (A, double point mutants; B, position +3 mutants) are indicated in Fig. 2.

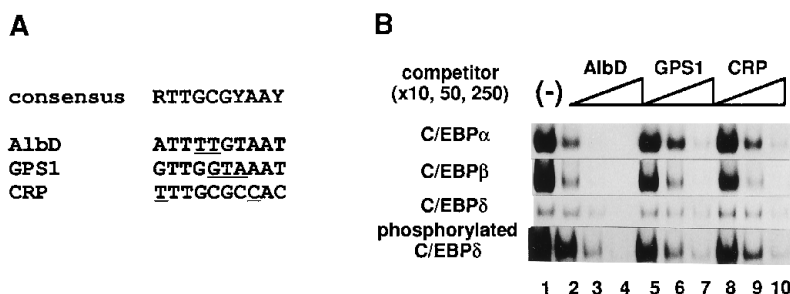


FIG. 5. Relative affinities of bacterially expressed C/EBP isoforms and phosphorylated C/EBP $\delta$  for the native binding sequences found in the genes regulated by the C/EBP family. A, shown are the nucleotide sequences of binding sites in the promoter of C/EBP regulated genes. The bases distinct from those of the consensus binding site are underlined. B, the probe (pal) containing palindrome consensus sequence was incubated with bacterially expressed proteins or phosphorylated C/EBP $\delta$  in the absence (lane 1) or presence (lanes 2–10) of increasing amounts of the competitor (10-, 50-, and 250-fold molar excesses) for gel shift analysis.

ing site located in chromatin structure, and it is possible that these interact synergistically with other transcription factors. As C/EBP $\beta$  interacts with NF- $\kappa$ B-p50 polypeptide and glucocorticoid receptor and C/EBP $\alpha$  associates with TFIIB and TATA-binding protein, the differential activity of C/EBP isoforms *in vivo* may be caused by distinct interactions with other proteins (33–35). *In vivo* experiments will be required to clarify the functional specificity of the C/EBP transcription factor family.

**Acknowledgments**—We thank Dr. Steven L. McKnight (Tularic, Inc., San Francisco, CA) for kindly providing C/EBP $\alpha$  cDNA. We are also grateful to Dr. John Goodier (National Institutes of Health) for critical reading of the manuscript.

## REFERENCES

1. Landschulz, W. H., Johnson, P. F., Adashi, E. Y., Graves, B. J., and McKnight, S. L. (1988) *Genes & Dev.* **2**, 786–800
2. Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishino, Y., Nakajima, T., Hirano, T., and Kishimoto, T. (1990) *EMBO J.* **9**, 1897–1906
3. Descombes, P., Chojkier, M., Lichtsteiner, S., Falvey, E., and Schibler, U. (1990) *Genes & Dev.* **4**, 1541–1551
4. Poli, V., Mancini, F. P., and Cortese, R. (1990) *Cell* **63**, 643–653
5. Imagawa, M., Osada, S., Koyama, Y., Suzuki, T., Hirom, P. C., Diccianni, M. B., Morimura, S., and Muramatsu, M. (1991) *Biochem. Biophys. Res. Commun.* **179**, 293–300
6. Williams, S. C., Cantwell, C. A., and Johnson, P. F. (1991) *Genes & Dev.* **5**, 1553–1567
7. Cao, Z., Umek, R. M., and McKnight, S. L. (1991) *Genes & Dev.* **5**, 1538–1552
8. Kinoshita, S., Akira, S., and Kishimoto, T. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 1473–1476
9. Kageyama, R., Sasai, Y., and Nakanishi, S. (1991) *J. Biol. Chem.* **266**, 15525–15531
10. Roman, C., Platero, J. S., Shuman, J., and Calame, K. (1990) *Genes & Dev.* **4**, 1404–1415
11. Nishizawa, M., Wakabayashi-Ito, N., and Nagata, S. (1991) *FEBS Lett.* **282**, 95–97
12. Ron, D., and Habener, J. F. (1992) *Genes & Dev.* **6**, 439–453
13. Ryden, T. H., and Beemon, K. (1989) *Mol. Cell. Biol.* **9**, 1155–1164
14. Merika, M., and Orkin, S. H. (1993) *Mol. Cell. Biol.* **13**, 3999–4010
15. Ko, L. J., and Engel, J. D. (1993) *Mol. Cell. Biol.* **13**, 4011–4022
16. Tanaka, N., Kawakami, T., and Taniguchi, T. (1993) *Mol. Cell. Biol.* **13**, 4531–4538
17. Sessa, G., Morelli, G., and Ruberti, I. (1993) *EMBO J.* **12**, 3507–3517
18. Benbrook, D. M., and Jones, N. C. (1994) *Nucleic Acids Res.* **22**, 1463–1469
19. Nishikawa, J., Kitaura, M., Matsumoto, M., Imagawa, M., and Nishihara, T. (1994) *Nucleic Acids Res.* **22**, 2902–2907
20. Mahoney, C. W., Shuman, J., McKnight, S. L., Chen, H.-C., and Huang, K.-P. (1992) *J. Biol. Chem.* **267**, 19396–19403
21. Trautwein, C., Van der Geen, P., Karin, M., Hunter, T., and Chojkier, M. (1994) *J. Clin. Invest.* **93**, 2554–2561
22. Ray, A., and Ray, B. K. (1994) *Mol. Cell. Biol.* **14**, 4324–4332
23. Friedman, A. D., Landschulz, W. H., and McKnight, S. L. (1989) *Genes & Dev.* **3**, 1314–1322
24. Inoue, S., Kondo, S., Hashimoto, M., Kondo, T., and Muramatsu, M. (1991) *Nucleic Acids Res.* **19**, 4091–4096
25. Hattori, M., and Sakai, Y. (1986) *Anal. Biochem.* **152**, 232–238
26. Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989) *Science* **246**, 911–916
27. Manak, J. R., de Bisschop, N., Kris, R. M., and Prywes, R. (1990) *Genes & Dev.* **4**, 955–967
28. Lichtsteiner, S., Wuarin, J., and Schibler, U. (1987) *Cell* **51**, 963–973
29. Imagawa, M., Osada, S., Okuda, A., and Muramatsu, M. (1991) *Nucleic Acids Res.* **19**, 5–10
30. Majello, B., Arcone, R., Toniatti, C., and Ciliberto, G. (1990) *EMBO J.* **9**, 457–465
31. Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. C. (1992) *Cell* **71**, 1223–1237
32. Krebs, E. G., Eisenman, R. N., Kuenzel, E. A., Litchfield, D. W., Lozeman, F. J., Luscher, B., and Sommercorn, J. (1988) *Cold Spring Harbor Symp. Quant. Biol.* **53**, 77–84
33. LeClair, K. P., Blarar, M. A., and Sharp, P. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8145–8149
34. Nshio, Y., Isshiki, H., Kishimoto, T., and Akira, S. (1993) *Mol. Cell. Biol.* **13**, 1854–1862
35. Nerlov, C., and Ziff, E. B. (1995) *EMBO J.* **14**, 4318–4328