ELSEVIER

Gene 196 (1997) 219–229

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

The chicken CCAAT/Enhancer Binding Protein α gene. Cloning, characterisation and tissue distribution

Cor. F. Calkhoven [1], Sonja I. Gringhuis, Geert AB *

Laboratory of Biochemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Received 16 December 1996; accepted 26 March 1997; Received by D. Schlessinger

Abstract

We present the cloning and sequencing of the gene encoding the chicken CCAAT/Enhancer Binding Protein α (cC/EBPα). The coding region and 1.5 kb of 5′ flanking DNA form a CpG island. Comparison of the chicken C/EBPα sequence to the homologous proteins of other species reveals several evolutionary conserved regions. cC/EBPα mRNA expression is restricted to a subset of tissues with high expression in liver, lung and small intestine. Recombinant cC/EBPα binds to its cognate C/EBP binding site as a homodimer or as a heterodimer with the related cC/EBPβ/NF-M. © 1997 Elsevier Science B.V.

Keywords: CpG-island; mRNA tissue distribution; Protein homology; DNA-binding

1. Introduction

The CCAAT/Enhancer Binding Protein α (C/EBPα) is a heat-stable DNA-binding protein that was first isolated and cloned from rat liver (Graves et al., 1986; Johnson et al., 1987). Its cognate DNA-binding site with the consensus sequence $T^{T}/_{G}NNGNAA^{T}/_{G}$ (Ryden and Beemon, 1989; Akira et al., 1990) is found in promoters of tissue-specific genes, including those expressed in hepatocytes, adipocytes and myelocytes (Christy et al., 1989; Friedman et al., 1989; Kaestner et al., 1990; McKnight, 1992; Ness et al., 1993). C/EBPα and other members of the C/EBP family show cell-type specific temporal expression patterns during cell differentiation. In the liver of developing mice, for example, C/EBPα mRNA is first detected just before birth, with a peak accumulation occurring at birth and lower levels in suckling pups and adults (Birkenmeier et al., 1989). C/EBPα induces growth arrest in a variety of cell types (Umek et al., 1991; Hendricks-Taylor and Darlington, 1995), and it appears that quiescent, terminally differentiated cells are the major sites of C/EBPα protein accumulation (Birkenmeier et al., 1989; McKnight, 1992). Earlier studies (see McKnight et al., 1989; Darlington et al., 1995) and a recent study with C/EBPα knockout mice (Wang et al., 1995) showed that C/EBPα is a critical regulator of genes governing integrative metabolic processes.

C/EBP proteins form a subclass of the bZIP (basic region leucine zipper) proteins. bZIP proteins are characterized by a DNA-binding domain consisting of a leucine zipper dimerisation region and a positively charged DNA-contacting region (Vinson et al., 1989). At present, six members of the C/EBP family designated C/EBPα-ζ have been cloned from various species (Landschulz et al., 1988; Xanthopoulos et al., 1989; Akira et al., 1990; Chang et al., 1990; Descombes et al., 1990; Poli et al., 1990; Roman et al., 1990; Cao et al., 1991; Akira et al., 1990; Williams et al., 1991; Kagayama et al., 1991; Calkhoven et al., 1992; Kinoshita et al., 1992; Ron and Habener, 1992; Thomassin et al., 1992; Xu and Tata, 1992; Katz et al., 1993; Chen et al., 1994; Antonson and Xanthopoulos, 1995). Whereas all C/EBPs bind DNA in a similar fashion to C/EBPα, they exhibit different effects in gene regulation. The complexity of the C/EBP family is further expanded by the translation of each C/EBPα and C/EBPβ mRNA into different isoforms, a transcriptionally active full-length isoform and a transcriptionally less active N-terminal

* Corresponding author. Tel: +31 50 3634182; Fax: +31 50-3634165; e-mail: G.AB@chem.rug.nl

[1] Present address: Max Delbrück-Centre for Molecular Medicine, Robert Rössle Strasse 10, 13125 Berlin, Germany.

Abbreviations: bp, base pair; bZIP, basic-zipper; C/EBP, CCAAT/enhancer binding protein; CR, conserved region; EMSA, electro mobility shift assay; NF-M, nuclear factor-myeloid; mRNA, (u)ORF, (upstream) open reading frame.

0378-1119/97/$17.00 © 1997 Elsevier Science B.V. All rights reserved.
PII S0378-1119(97)00230-8

truncated isoform initiated from an internal, in-frame translation start codon (Descombes and Schibler, 1991; Calkhoven et al., 1992; Lin et al., 1993; Ossipow et al., 1993).

2. Materials and methods

2.1. Cloning of the chicken C/EBPα gene

A chicken genomic DNA library in EMBL3 (Clontech) was screened with a partial cC/EBP cDNA probe of 480 bp (Wijnholds et al., 1991). *E. coli* strain NM539, infected with 2×10^6 pfu, was plated on to 22×22-cm Nunc plates with NZYM medium (Sambrook et al., 1989). Replicates of the plaques on Nytran-N were made in duplicate. The DNA on the membrane was denatured and immobilized by baking, as described by the manufacturer (Schleicher & Schuell). Hybridisation and washing steps were carried out as described by Church and Gilbert (1984). Blots were prehybridized at 65°C for 2 h and hybridized overnight at 65°C with 100 ng [α-^{32}P]dCTP-labelled probe (random primed labelling kit, Boehringer) in hybridisation solution [0.5 M $NaHPO_4$; 1% bovine serum albumin (BSA); 1 mM Na_2EDTA; 7% SDS]. The membranes were washed at 65°C, once in washing solution [40 mM $NaHPO_4$; 1 mM Na_2EDTA] with 5% SDS and five to seven times in washing solution with 1% SDS. The membranes were exposed to Kodak X-ray film. Plaques corresponding to duplicate signals were plugged out, used for reinfection, and replated to purity.

2.2. Subcloning and sequencing of cC/EBPα

From a 6.9-kb *Eco*RI fragment contained in the genomic clones, three *Sac*I fragments covering the coding region of cC/EBPα were subcloned in pGEM7(+) (Promega) and sequenced. A 1.85-kb *Pst*I fragment covering the upstream flanking sequences was cloned and sequenced by subsequent subcloning of smaller restriction fragments and by using internal primers. Sequencing of denatured double-stranded DNA was performed following the method of Sanger et al. (1977) using T7 DNA polymerase (Pharmacia) and Taq DNA polymerase (Promega) with 7-deaza-dGTP substituting dGTP to reduce band compression. Both strands were sequenced independently.

2.3. RNA analysis

Total RNA was isolated from tissues homogenized in guanidinium thiocyanate denaturing solution as described by Ausubel et al. (1993). Northern blotting on Nytran-N membranes was performed as described by the manufacturer (Schleicher & Schuell). Hybridisation solutions were as described by Church and Gilbert (1984). Blots were prehybridized at 65°C for 2 h and hybridized overnight at 65°C with 500 000 cpm/ml [α-^{32}P]dCTP-labelled probe. The membranes were washed at 65°C once with 5% SDS washing solution and five to seven times with 1% SDS washing solution. Northern blots were hybridized with a ^{32}P-labelled 300-bp cC/EBPα probe excised from cC/EBP:ΔdomI-pSG5 (Calkhoven et al., 1994) with *Bam*HI and *Eco*RI. As a control on RNA integrity, blots were probed with a 650-bp fragment from the chicken β-actin cDNA. The probes were [α-^{32}P]dCTP-labelled with the random-primed labelling kit (Boehringer-Mannheim). Membranes were exposed to Kodak X-ray film.

2.4. Sequence analysis

The CpG island was identified using the criteria of Gardiner-Garden and Frommer (1987). The ratio observed/expected CpG dinucleotides was calculated as follows:

$$\text{Obs/Exp CpG} = (\text{number of CpG/number of C} \times \text{number of G}) \times \text{N} \; (= \text{total number of nucleotides}).$$

CpG-rich regions are defined as DNA stretches with both an average G+C ≥50% and an Obs/Exp CpG greater than 0.6. The graph of the G+C/ATGC-ratio was made using the program PBASE of PCgene (IntelliGenetics Inc.). For alignment of C/EBPα amino acid sequences and the construction of the phylogenetic tree, the MegAlign program was used (DNASTAR, Inc.).

2.5. Construction of expression plasmids and overexpression of recombinant C/EBP proteins in Escherichia coli

Restriction sites for subcloning in pET prokaryotic expression vectors (Studier et al., 1990) were created by site-directed mutagenesis of the cC/EBPwt-pSG5 construct as previously described (Calkhoven et al., 1994). *cC/EBPα-pET3c*: A *Bam*HI site was introduced just in front of the first ATG start codon using primer 5′-AGG CTG TAG GAT CCC CAT GGA GCA AGC-3′ (mutations underlined). The *Bam*HI fragment covering the cC/EBPα coding region was cloned in the *Bam*HI site of pET3c, and the obtained clones were selected by restriction analysis for the proper orientation. For *bZIPα-pET3b*, a *Sma*I–*Bam*HI fragment of cC/EBPwt-pSG5 covering the bZIP region of cC/EBPα was blunted and cloned in the *Sma*I site of pBluescript KSII to introduce a *Bam*HI site just upstream from the *Sma*I site. The *Bam*HI fragment from the pBluescript construct was cloned in pET3a, and the obtained clones

were selected by restriction analysis for the proper orientation. For *cC/EBPβ/NF-M-pET3d*, an *Nco* I–*Bam*HI fragment from the NF-M-pBluescript (a gift of A. Leutz, MDC Berlin), containing the coding region of NF-M, was cloned in the *Nco*I–*Bam*HI sites of pET3d.

For overexpression in *E. coli,* cC/EBPα-pETc, bZIPα-pET3b and NF-M.pETd constructs were transfected into strain BL21DE3.pLysE (Studier et al., 1990). Cultures were grown at 28°C to $OD_{660}=0.6$ and induced with 1 mM IPTG. After 4 h at 28°C, cells were harvested, washed once in buffer and resuspended in 1/100 volume buffer [100 mM KCl, 25 mM HEPES (pH 7.8), 12.5 mM $MgCl_2$, 1 mM DTT, 20% glycerol]. Total cellular protein extracts were made by sonication and removal of precipitated material by 30-min centrifugation in an Eppendorf centrifuge. The soluble protein fraction contained 10–30% recombinant overexpressed protein. Because C/EBP proteins are heat-stable, the recombinant cC/EBP proteins could be partially purified by incubating the total protein extract for 10 min at 70°C and removal of the precipitated material by 3-min centrifugation in an Eppendorf centrifuge.

2.6. Electro mobility shift assays (EMSA)

Probe 'D' used in the EMSA was obtained by annealing of the synthetic HPLC purified oligonucleotides: 5′-GGG ACC TTC CAT TAC CAA ATC CGA ACC C-3′ and 5′-GGG TTC GGA TTT GGT AAT GGA AGG TCC C-3′: 50 ng of one of the single-stranded oligonucleotides was ^{32}P-labelled with T4-polynucleotide kinase. Labelled DNA was separated from free ^{32}P-label by gel filtration over Sephadex G50, precipitation with ethanol in the presence of 2 μg poly(dIdC) and washing with 96% ethanol. The labelled DNA was dissolved with a fourfold excess of the opposite DNA strand in 20 μl annealing buffer [10 mM Tris–HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT]. Prior to annealing of the strands, the mixture was heated for 2 min at 90°C and cooled slowly to room temperature.

The binding reaction in a final volume of 15 μl contained binding buffer [20 mM HEPES (pH 7.9), 60 mM KCl, 0.06 mM EDTA, 0.6 or 2.5 mM DTT, 2 mM spermidine, 10% glycerol], supplemented with 2 μg poly(dIdC), 0.1–0.5 ng labelled double-stranded oligonucleotides. The binding reaction containing 1–3 μg bacterial extract was preincubated for 10 min on ice before adding the DNA probe, followed by a 20-min incubation at 25°C. The samples were analysed on a 4% (30:1) PAA gels with 0.5 TBE as running buffer. After air-drying on Whatman paper, the gels were exposed to Kodak X-ray film.

3. Results

3.1. The nucleotide sequence and deduced amino acid sequence of cC/EBPα

Screening of a genomic chicken DNA library in λEMBL3 (Clonetech) with a C/EBPα cDNA probe yielded three overlapping clones. The coding region (EMBL accession no: X66844) and ~1.5 kb of 5′ flanking DNA were sequenced (Fig. 1). Like the C/EBPαs from other species, the chicken C/EBPα gene is intronless and contains an uninterrupted reading frame of 971 bp. In the flanking sequences, a TATA-box and a CAP-signal are present at 242 bp and 212 bp 5′ to the C/EBPα ORF, respectively, and two AATAAA polyadenylation motifs at 133 bp and 302 bp 3′ to the ORF. The cC/EBPα ORF encodes a polypeptide of 324 amino acids with, at its C-terminus, the bipartite bZIP domain (Fig. 1B). Just upstream of the cC/EBPα ORF lies a small ORF, potentially encoding a pentapeptide that is conserved with respect to length and position in rat, mouse, chicken, *Xenopus*, frog and human. Its presence in all vertebrate C/EBPs sequenced to date emphasises that the uORF must have a significant physiological function. We have shown earlier that the uORF is crucial for the formation of a truncated C/EBPα isoform from an internal in-frame AUG-codon (Calkhoven et al., 1994).

A 2.2-kb gene region, including the larger part (1250 bp) of the chicken C/EBPα transcribed sequence and 950 bp of upstream sequence, has a G+C content of 73.5%, markedly higher than the 48 and 42% G+C content of the sequences flanking to the 5′ and 3′ side, respectively (Fig. 1A). This indicates that the region may represent a CpG island. To be definitively defined as a CpG island, the sequence should conform to a second criterion reading that the observed/expected CpG-dinucleotide frequency is larger than 0.6 (Gardiner-Garden and Frommer, 1987). For the average vertebrate DNA, except CpG islands, the CpG-dinucleotide frequency is much lower because of deamination of 5-methylcytosine to thymine (Coulondre et al., 1978; Gardiner-Garden and Frommer, 1987). The CG-rich region of the chicken C/EBPα gene locus has a CpG-dinucleotide frequency close to the calculated value (CpG obs/exp = 0.97) showing that its cytosines have been kept unmethylated in cells of the germ line and defining it as a classical CpG island.

C/EBPα is known to be subject to autoregulation either by direct binding of the C/EBP protein to the upstream regulatory element as was shown for the murine C/EBPα gene (Christy et al., 1991; Legraverend et al., 1993; Rana et al., 1995) or by indirect positive feedback via an USF binding site as was shown for the human C/EBPα gene (Timchenko et al., 1995). In this respect, it is interesting that chicken C/EBPα contains three putative C/EBP binding sites in the upstream sequences of the cC/EBPα gene. One site is located in

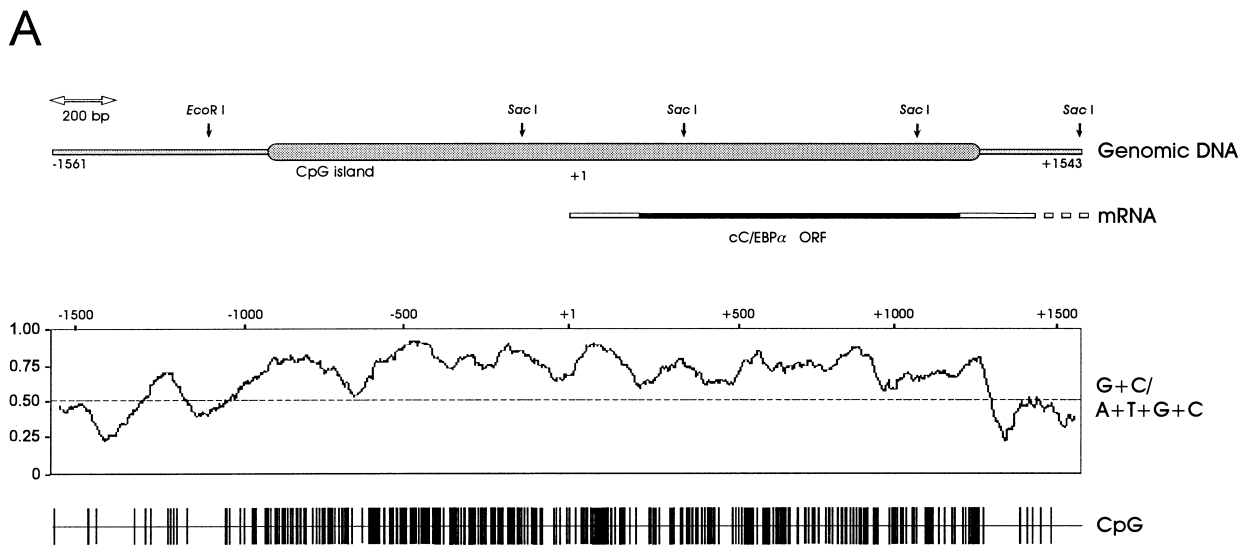


Fig. 1. Physical map of the genomic cC/EBPα gene region. (**A**) The map covers the sequenced region of the cC/EBPα gene with the position of the CpG island; the transcribed sequences and the position of the cC/EBPα ORF indicated. Restriction sites used for subcloning are also indicated. The lower graph displays the G+C/A+T+G+C ratio as calculated for a 75-nt window (PCgene, IntelliGenetics, Inc.) and the distribution of CpG dinucleotides indicated by vertical lines. (**B**) Nucleotide sequence and deduced amino acid sequence of cC/EBPα. The nucleotide sequence from −1560 to +1544 is shown. The TATA-box and the putative CAP-site are in bold and double-underlined; the AATAAA- polyadenylation motifs are also double-underlined. Translation initiation codons used in vivo are in bold, including an upstream CTG start site, with the surrounding sequences important for translation initiation efficiency underlined (Kozak, 1989; Boeck and Kalakofsky, 1994; Calkhoven et al., 1994; Grünert and Jackson, 1994). From the M3-ATG start site, a smaller isoform is generated (Calkhoven et al., 1994). The leucines in the leucine zipper dimerisation interface are indicated in bold italics. The deduced pentapeptide encoded by the small uORF is shown in italics. The putative C/EBPα recognition sites in the upstream flanking region are highlighted in a shaded box.

the CpG island around position −642 (5′-TGGGGAAAT-3′) and two sites outside the CpG island, around −1412 (5′-TTCTGAAAT-3′ non-coding strand) and −1435 (5′-TGCTGAAa-3′) (mismatch with consensus 5′-$T^{T}/_{G}$NNGNAA$^{T}/_{G}$-3′ in lower case) (Fig. 1B, boxed). The presence of these C/EBP-binding sites suggests that chicken C/EBPα, like the murine gene, may be subject to direct autoregulation.

3.2. Conserved regions in the cC/EBPα protein

Comparison of the C/EBPα proteins of rat (Landschulz et al., 1988; Lincoln et al., 1994), chicken, *Xenopus* (Xu and Tata, 1992), frog (*Rana catesbeiana*) (Chen et al., 1994) and human (Antonson and Xanthopoulos, 1995) reveals several evolutionary conserved regions (Fig. 2). The C-terminal DNA-binding domain consisting of the basic DNA-contacting and leucine zipper dimerisation regions is virtually identical for all these species. Besides the bZIP domains, four conserved regions named N-term, CRI, CRII and CRIII can be distinguished, which most likely represent functional elements mediating specific functions. Sequences in the region covering N-term and CRI are responsible for the major *trans*activation function as was shown in several studies (Pei and Shih, 1991; Ossipow et al., 1993; Calkhoven et al., 1994; Nerlov and Ziff, 1994, 1995). This region is lacking in the smaller translational isoform of C/EBPα, which, depending on the promoter context, acts as a less active or negative transcriptional regulator in liver cells (Ossipow et al., 1993; Calkhoven et al., 1994; Nerlov and Ziff, 1994). The relationship between the different C/EBPα sequences is shown in a dendogram (Fig. 2B).

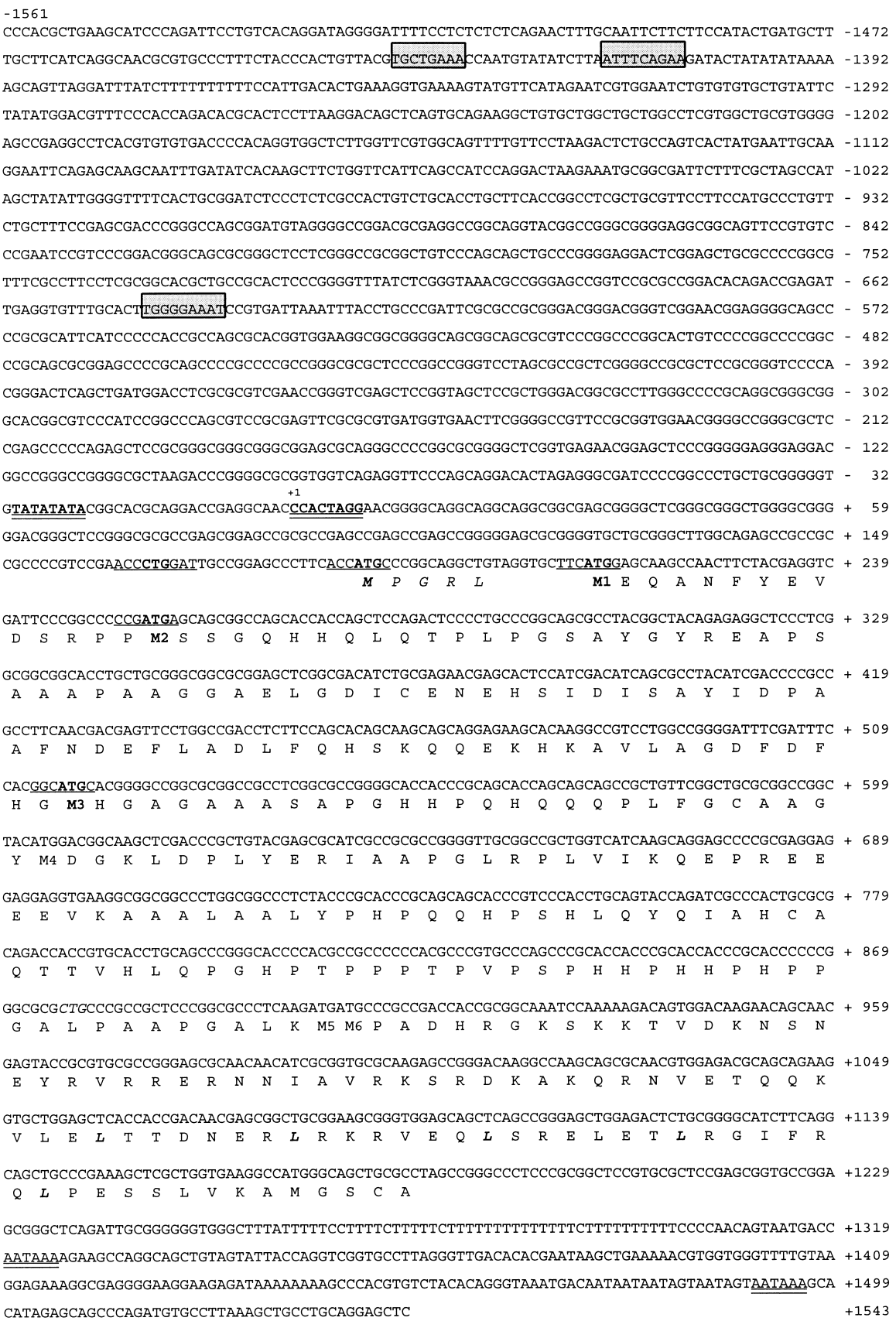
3.3. Tissue distribution of cC/EBPα mRNA

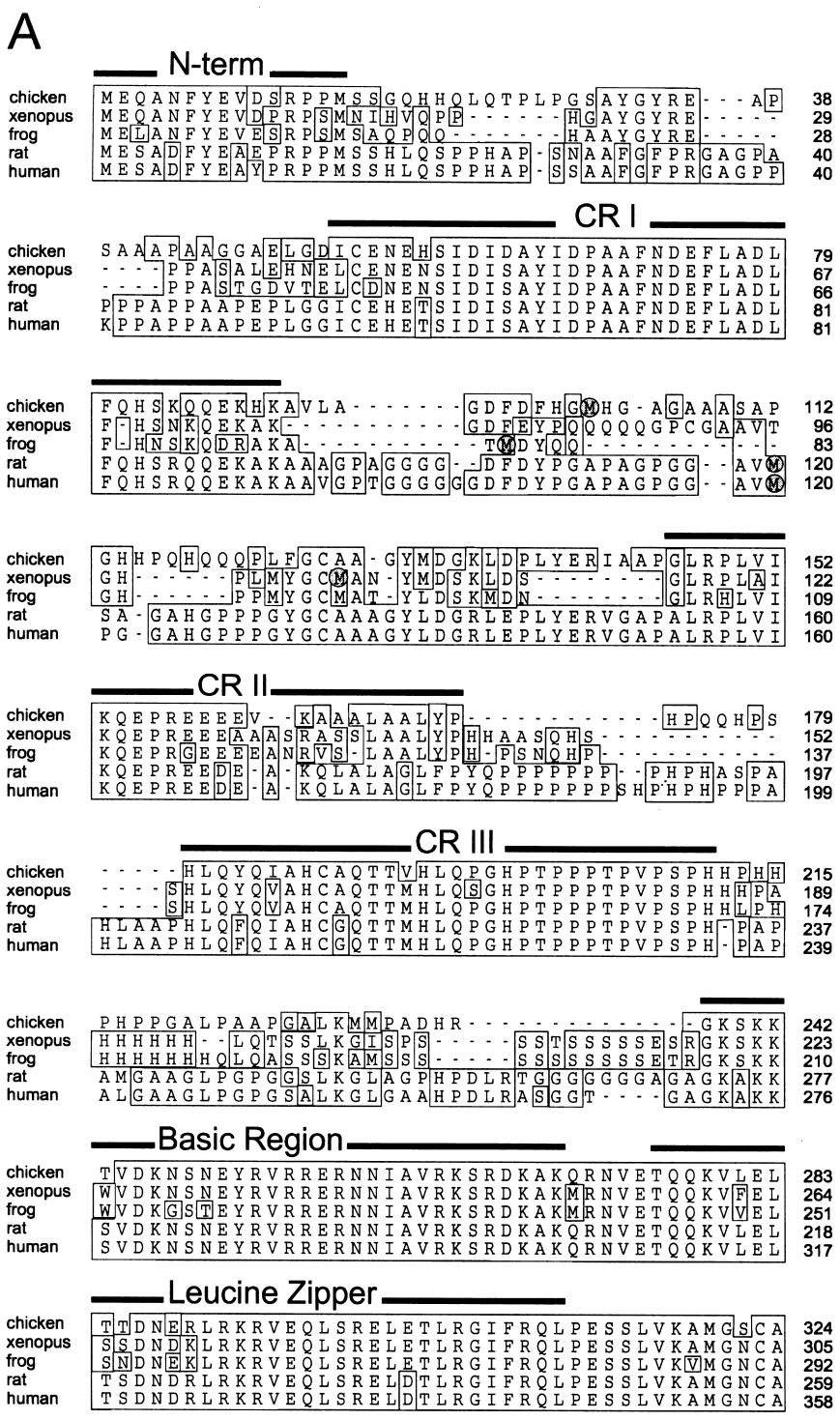
The expression of cC/EBPα mRNA in various chicken tissues was examined by Northern blotting. The highest levels of cC/EBPα mRNA in chicken are found in liver, lung and small intestine. Lower amounts are detectable in duodenum, spleen and kidney (Fig. 3). Comparison of our Northern analysis with those published for murine, human, *Xenopus* and frog (Table 1) shows that high levels of C/EBPα mRNA are found in the liver of all species (Williams et al., 1991; Xu and Tata, 1992; Chen et al., 1994; Antonson and Xanthopoulos, 1995). Notable differences between the five species are found, especially in kidney, skeletal muscle and heart. Whether these reflect species-specific differences in metabolic functions is not known.

3.4. Recombinant protein isoforms of cC/EBPα and cC/EBPβ/NF-M mutually form DNA-binding dimers

Chicken C/EBPβ/NF-M is a member of the C/EBP-family (Katz et al., 1993), which is expressed abundantly

B


-1561
CCCACGCTGAAGCATCCCAGATTCCTGTCACAGGATAGGGGATTTTCCTCTCTCTCAGAACTTTGCAATTCTTCTTCCATACTGATGCTT -1472
TGCTTCATCAGGCAACGCGTGCCCTTTCTACCCACTGTTACGTGCTGAAACCAATGTATATCTTAATTTCAGAAGATACTATATATAAAA -1392
AGCAGTTAGGATTTATCTTTTTTTTTTCCATTGACACTGAAAGGTGAAAAGTATGTTCATAGAATCGTGGAATCTGTGTGTGCTGTATTC -1292
TATATGGACGTTTCCCACCAGACACGCACTCCTTAAGGACAGCTCAGTGCAGAAGGCTGTGCTGGCTGCTGGCCTCGTGGCTGCGTGGGG -1202
AGCCGAGGCCTCACGTGTGTGACCCCACAGGTGGCTCTTGGTTCGTGGCAGTTTTGTTCCTAAGACTCTGCCAGTCACTATGAATTGCAA -1112
GGAATTCAGAGCAAGCAATTTGATATCACAAGCTTCTGGTTCATTCAGCCATCCAGGACTAAGAAATGCGGCGATTCTTTCGCTAGCCAT -1022
AGCTATATTGGGGTTTTCACTGCGGATCTCCCTCTCGCCACTGTCTGCACCTGCTTCACCGGCCTCGCTGCGTTCCTTCCATGCCCTGTT - 932
CTGCTTTCCGAGCGACCCGGGCCAGCGGATGTAGGGGCCGGACGCGAGGCCGGCAGGTACGGCCGGGCGGGGAGGCGGCAGTTCCGTGTC - 842
CCGAATCCGTCCCGGACGGGCAGCGCGGGCTCCTCGGGCCGCGGCTGTCCCAGCAGCTGCCCGGGGAGGACTCGGAGCTGCGCCCCGGCG - 752
TTTCGCCTTCCTCGCGGCACGCTGCCGCACTCCCGGGGTTTATCTCGGGTAAACGCCGGGAGCCGGTCCGCGCCGGACACAGACCGAGAT - 662
TGAGGTGTTTGCACTTGGGGAAATCCGTGATTAAATTTACCTGCCCGATTCGCGCCGCGGGACGGGACGGGTCGGAACGGAGGGGCAGCC - 572
CCGCGCATTCATCCCCCACCGCCAGCGCACGGTGGAAGGCGGCGGGGCAGCGGCAGCGCGTCCCGGCCCGGCACTGTCCCCGGCCCCGGC - 482
CCGCAGCGCGGAGCCCCGCAGCCCCGCCCCGCCGGGCGCGCTCCCGGCCGGGTCCTAGCGCCGCTCGGGGCCGCGCTCCGCGGGTCCCCA - 392
CGGGACTCAGCTGATGGACCTCGCGCGTCGAACCGGGTCGAGCTCCGGTAGCTCCGCTGGGACGGCGCCTTGGGCCCCGCAGGCGGGCGG - 302
GCACGGCGTCCCATCCGGCCCAGCGTCCGCGAGTTCGCGCGTGATGGTGAACTTCGGGGCCGTTCCGCGGTGGAACGGGGCCGGGCGCTC - 212
CGAGCCCCCAGAGCTCCGCGGGCGGGCGGGCGGAGCGCAGGGCCCCGGCGCGGGGCTCGGTGAGAACGGAGCTCCCGGGGGAGGGAGGAC - 122
GGCCGGGCCGGGGCGCTAAGACCCGGGGCGCGGTGGTCAGAGGTTCCCAGCAGGACACTAGAGGGCGATCCCCGGCCCTGCTGCGGGGGT - 32
+1
GTATATATACGGCACGCAGGACCGAGGCAACCCACTAGGAACGGGGCAGGCAGGCAGGCGGCGAGCGGGGCTCGGGCGGGCTGGGGCGGG + 59
GGACGGGCTCCGGGCGCGCCGAGCGGAGCCGCGCCGAGCCGAGCCGAGCCGGGGGAGCGCGGGGTGCTGCGGGCTTGGCAGAGCCGCCGC + 149
CGCCCCGTCCGAACCCTGGATTGCCGGAGCCCTTCACCATGCCCGGCAGGCTGTAGGTGCTTCATGGAGCAAGCCAACTTCTACGAGGTC + 239
M P G R L M1 E Q A N F Y E V
GATTCCCGGCCCCCGGATGAGCAGCGGCCAGCACCACCAGCTCCAGACTCCCCTGCCCGGCAGCGCCTACGGCTACAGAGAGGCTCCCTCG + 329
D S R P P M2 S S G Q H H Q L Q T P L P G S A Y G Y R E A P S
GCGGCGGCACCTGCTGCGGGCGGCGCGGAGCTCGGCGACATCTGCGAGAACGAGCACTCCATCGACATCAGCGCCTACATCGACCCCGCC + 419
A A A P A A G G A E L G D I C E N E H S I D I S A Y I D P A
GCCTTCAACGACGAGTTCCTGGCCGACCTCTTCCAGCACAGCAAGCAGCAGGAGAAGCACAAGGCCGTCCTGGCCGGGGATTTCGATTTC + 509
A F N D E F L A D L F Q H S K Q Q E K H K A V L A G D F D F
CACGGCATGCACGGGGCCGGCGCGGCCGCCTCGGCGCCGGGGCACCACCCGCAGCACCAGCAGCAGCCGCTGTTCGGCTGCGCGGCCGGC + 599
H G M3 H G A G A A A S A P G H H P Q H Q Q Q P L F G C A A G
TACATGGACGGCAAGCTCGACCCGCTGTACGAGCGCATCGCCGCGCCGGGGGTTGCGGCCGCTGGTCATCAAGCAGGAGCCCCGCGAGGAG + 689
Y M4 D G K L D P L Y E R I A A P G L R P L V I K Q E P R E E
GAGGAGGTGAAGGCGGCGGCCCTGGCGGCCCTCTACCCGCACCCGCAGCAGCACCCGTCCCACCTGCAGTACCAGATCGCCCACTGCGCG + 779
E E V K A A A A L A A L Y P H P Q Q H P S H L Q Y Q I A H C A
CAGACCACCGTGCACCTGCAGCCCGGGCACCCCACGCCGCCCCCCACGCCCGTGCCCAGCCCGCACCACCCGCACCACCCGCACCCCCCG + 869
Q T T V H L Q P G H P T P P P T P V P S P H H P H H P H P P
GGCGCGCTGCCCGCCGCTCCCGGCGCCCTCAAGATGATGCCCGCCGACCACCGCGGCAAATCCAAAAAGACAGTGGACAAGAACAGCAAC + 959
G A L P A A P G A L K M5 M6 P A D H R G K S K K T V D K N S N
GAGTACCGCGTGCGCCGGGAGCGCAACAACATCGCGGTGCGCAAGAGCCGGGACAAGGCCAAGCAGCGCAACGTGGAGACGCAGCAGAAG +1049
E Y R V R R E R N N I A V R K S R D K A K Q R N V E T Q Q K
GTGCTGGAGCTCACCACCGACAACGAGCGGCTGCGGAAGCGGGTGGAGCAGCTCAGCCGGGAGCTGGAGACTCTGCGGGGCATCTTCAGG +1139
V L E L T T D N E R L R K R V E Q L S R E L E T L R G I F R
CAGCTGCCCGAAAGCTCGCTGGTGAAGGCCATGGGCAGCTGCGCCTAGCCGGGCCCTCCCGCGGCTCCGTGCGCTCCGAGCGGTGCCGGA +1229
Q L P E S S L V K A M G S C A
GCGGGCTCAGATTGCGGGGGGTGGGCTTTATTTTTCCTTTTCTTTTTCTTTTTTTTTTTTTCTTTTTTTTTTCCCCAACAGTAATGACC +1319
AATAAAAGAAGCCAGGCAGCTGTAGTATTACCAGGTCGGTGCCTTAGGGTTGACACACGAATAAGCTGAAAAACGTGGTGGGTTTTGTAA +1409
GGAGAAAGGCGAGGGGAAGGAAGAGATAAAAAAAAGCCCACGTGTCTACACAGGGTAAATGACAATAATAATAGTAATAGTAATAAAGCA +1499
CATAGAGCAGCCCAGATGTGCCTTAAAGCTGCCTGCAGGAGCTC +1543


A
N-term
CR I
CR II
CR III
Basic Region
Leucine Zipper
chicken
xenopus
frog
rat
human
B
Xenopus
frog
chicken
rat
human

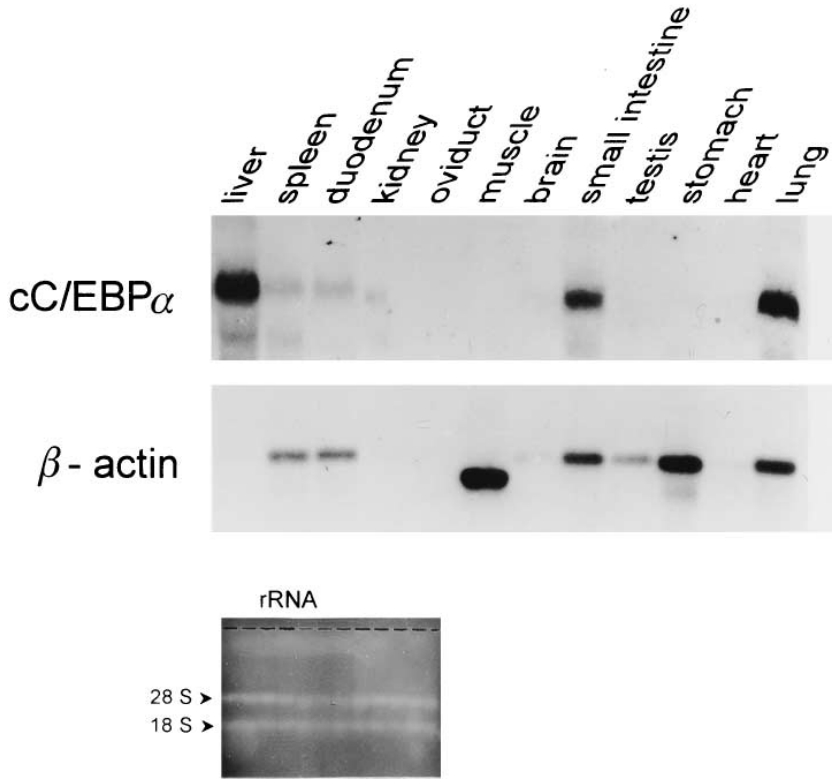


Fig. 3. Tissue distribution of cC/EBPα mRNA. 30 μg of total RNA was analysed by Northern analysis using DNA-probes of cC/EBPα and β-actin. The following tissues were used; from hen, spleen, duodenum, kidney, oviduct, skeletal muscle; and from rooster, liver, brain, small intestine, testis, stomach, heart, lung. The lower panel is a photograph of the ethidium bromide-stained gel showing the major rRNA bands.

in chicken liver and muscle as was shown by Northern analysis (Katz et al., 1993; Calkhoven et al., unpublished data). To test whether cC/EBPα and cC/EBPβ/NF-M are able to bind to DNA as homodimers and/or heterodimers, recombinant full-length C/EBP proteins alone or together with the bZIP domain of cC/EBPα where examined in an EMSA. Fig. 4 shows:

(1) that each of the individual proteins, full-length cC/EBP (α_{FL}), full-length cC/EBPβ (β_{FL})and bZIP-cC/EBPα (α_{bZIP}), bind to DNA and
(2) that mixtures of bZIP-cC/EBPα and either full-length cC/EBPα or full-length cC/EBPβ form DNA-binding complexes of intermediate mobility.

In conclusion, chicken C/EBPα and -β proteins readily form homo- and heterodimers.

4. Discussion

We have cloned and analysed the chicken CCAAT/enhancer binding protein, and the experiments presented in this article provide some clues of how expression of the cC/EBPα gene may be regulated at the transcriptional and post-translational level.

The sequence of the chicken C/EBPα gene region reveals that the cC/EBPα coding region and the contiguous upstream regulatory region comprise a CpG island. The methylation status of CpG islands is believed to play an important role in the regulation of gene expression during development and tissue differentiation (Becker et al., 1987; Li et al., 1992). Methylation of DNA has been shown to inhibit gene expression in vitro and in vivo (see Razin and Cedar, 1991; Eden and Cedar, 1994). Methyl groups may interfere with the binding of transcription factors to cognate target sites, thereby preventing the formation of an active transcription complex on the promoter (Comb and Goodman, 1990). Methylation of the CpG island of the myogenic transcription factor MyoD1 gene during the immortalisation of cells results in the down-regulation of MyoD1 expression (Jones et al., 1990). In this respect, it is noteworthy that C/EBPα is down-regulated in proliferat-

Fig. 2. Comparison of C/EBPα amino acid sequences. (**A**) The amino acid sequences of chicken, *Xenopus*, frog (*Rana catesbeiana*), rat and human C/EBPα are shown in their optimal alignment (MegAlign program, DNASTAR, Inc). Positional identities are boxed. The most conserved regions are overlined: N-term, CR I, CR II, CR III, basic region and leucine zipper. The encircled methionines lying between CRI and CRII are the N-terminal residues of the smaller C/EBP isoforms as established for murine and chicken (Descombes and Schibler, 1991; Calkhoven et al., 1994) or otherwise predicted from the sequence context of the ATG codon (Calkhoven et al., 1994). (**B**) The relationship between the different C/EBPα sequences is shown in the dendogram (MegAlign program, DNASTAR, Inc.).

Table 1
Comparison of the C/EBPα mRNA tissue distribution in chicken with other species

	Chicken[a]	Murine[b]	Human[c]	*Xenopus*[d]	Frog[e]
Liver	++++	+++	++++	++++	+++
Spleen	+	+	+	++	ND
Duodenum	++	ND	ND	ND	ND
Kidney	+	++	—	++	+++
Oviduct	—	ND	ND	++	ND
Skeletal muscle	—	+	+	—	—
Brain	—	+	+/-	+/-	—
Small intestine	+++	+	+++	ND	ND
Testis	—	+	—	ND	ND
Stomach	+/-	—	ND	ND	ND
Heart	—	+	++	—	—
Lung	++++	++	++	ND	ND

The relative mRNA levels can be compared only within the species.
[a]Total RNA.
[b]Data from Williams et al. (1991).
[c]Data from Antonson and Xanthopoulos (1995).
[d]Data from Xu and Tata (1992), total RNA.
[e]Data from Chen et al. (1994), total RNA.
ND, not determined.

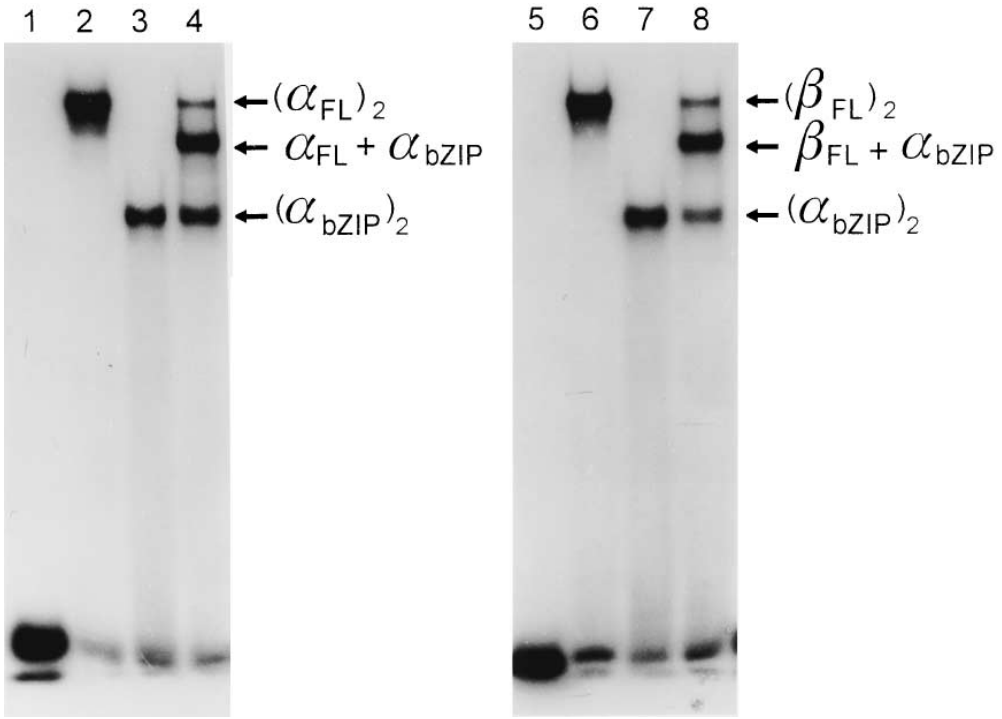


Fig. 4. DNA-binding of recombinant chicken C/EBPα and -β homo- and heterodimers. EMSA of recombinant cC/EBP proteins with the apoVLDL II-promoter C/EBP-binding site 'D' (Wijnholds et al., 1991). Lanes 1 and 5, free oligo; lane 2, full-length cC/EBPα (1 μg); lanes 3 and 7, bZIPα (3 μg); lane 4, mixture of bZIPα and cC/EBPα (3 μg and 1 μg); lane 6, full-length cC/EBPβ/NF-M (3 μg); lane 8, mixture of bZIPα and cC/EBPβ/NF-M (3 μg).

ing cells, e.g. in the human hepatoma cell line HepG2 (Friedman et al., 1989) and the chicken hepatoma cell line LMH (Calkhoven et al., unpublished data). Moreover, rat C/EBPα has been found to be downregulated during liver carcinogenesis, probably contributing to the proliferative advantage of neoplastic nodules (Flodby et al., 1995). Possibly, de novo methylation of CpG dinucleotides could be responsible for the downregulation of C/EBPα during immortalisation of liver cells.

Both rat and human C/EBPα have been shown to be subject to autoregulation. The mechanism appears to be different, depending on the species; in rat, it involves

direct binding of C/EBP to its own gene promoter (Christy et al., 1991; Legraverend et al., 1993; Rana et al., 1995) and in human, indirect action via the transcription factor USF (Timchenko et al., 1995). The three putative C/EBP binding sites just upstream from the CpG island of the chicken gene strongly suggest the action of a direct positive autoregulatory feedback mechanism in the maintenance of cC/EBPα expression. Unfortunately, homologous C/EBPα genes have, to our knowledge (DNA database), only been sequenced up to 450 bp upstream, thus far preventing a comparative sequence analysis of the respective regions to examine the possible generality of the upstream C/EBP sites.

Heterodimerisation between different members of the C/EBP family, as shown here for chicken C/EBPα and NF/M (=C/EBPβ), is expected to have implications for their action. Although the different C/EBPs may show some redundancy in function, they probably perform different regulatory roles. Whereas C/EBPα acts in the regulation of energy homeostasis (McKnight et al., 1989; Darlington et al., 1995), C/EBPβ appears to be implicated in the acute phase response, cytokine expression and inflammation (Akira et al., 1990; Poli et al., 1990). Moreover, the antimitotic activity appears to be unique for C/EBPα. In situations where the relative abundance of C/EBPβ increases over that of C/EBPα, as for example after partial hepatectomy (Diehl and Yang, 1994; Rana et al., 1995), the formation of C/EBPβ homodimers will be favoured over that of C/EBPα homo- and heterodimers. This might facilitate the entrance of hepatocytes into the proliferative phase associated with liver regeneration. Conversely, the ratio of C/EBPα:C/EBPβ drastically increases during the termination of clonal expansion of preadipocytes preceding the terminally differentiated state (Cao et al., 1991). In general, the homo- and heterodimer interactions between a limited set of transcription factors generates greater diversity and adds an additional level of complexity to differential gene expression.

Acknowledgement

We thank Hans Eshuis for assistance with sequencing, Harm Nijveen for preparing the cC/EBPα-pET prokaryotic expression vector, Justin Heydenrijk for EMSA and Achim Leutz (Max Delbrück-Centre, Berlin) for providing the NF-M/cC/EBPβ clone and NF-M antiserum. This work was supported by the Netherlands Organisation for Scientific Research (NWO) through the Council of Medical Research (GB-MW).

References

Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., Nakajima, T., Hirano, T., Kishimoto, T., 1990. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. EMBO J. 9, 1897–1906.

Antonson, P., Xanthopoulos, G., 1995. Molecular cloning, sequencing and expression patterns of the human gene encoding CCAAT/enhancer binding protein α (C/EBPα). Biochem. Biophys. Res. Commun. 215, 106–113.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 1993. Current Protocols in Molecular Biology. Green Publishing Associates and John Wiley and Sons, Brooklyn, NY.

Becker, P.B., Ruppert, S., Schutz, G., 1987. Genomic footprinting reveals cell type-specific DNA binding ubiquitous factors. Cell 51, 435–443.

Birkenmeier, E.H., Gwynn, B., Howard, S., Jerry, J., Gordon, J.L., Landschulz, W.H., McKnight, S.L., 1989. Tissue-specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein. Genes Dev. 3, 1146–1156.

Boeck, R., Kalakofsky, D., 1994. Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. EMBO J. 13, 3608–3617.

Calkhoven, C.F., AB, G., Wijnholds, J., 1992. cC/EBP, a chicken transcription factor of the leucine-zipper C/EBP family. Nucl. Acids Res. 20, 4093

Calkhoven, C.F., Bouwman, P.R.J., Snippe, L., AB, G., 1994. Translation start site multiplicity of the CCAAT/enhancer binding protein α mRNA is dictated by a small 5′ open reading frame. Nucl. Acids Res. 22, 5540–5547.

Cao, Z., Umek, R.M., McKnight, S.L., 1991. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. 5, 1538–1552.

Chang, C.-J., Chen, T.-T., Lei, H.-Y., Chen, D.-S., Lee, S.-C., 1990. Molecular cloning of a transcription factor, AGP/EBP, that belongs to members of the C/EBP family. Mol. Cell. Biol. 10, 6642–6653.

Chen, Y., Hu, H., Atkinson, B.G., 1994. Characterization and expression of C/EBP-like genes in the liver of *Rana catesbeiana* tadpoles during spontaneous and thyroid hormone-induced metamorphosis. Dev. Genet. 15, 366–377.

Christy, R.J., Yang, V., Ntambi, J., Geiman, D., Landschulz, W., Friedman, A., Nakabeppu, Y., Kelly, T., Lane, M.D., 1989. Differentiation-induced gene expression in 3T3-L1 preadipocytes: CCAAT/enhancer binding proteins interacts with and activates the promoters of two adipocyte-specific genes. Genes Dev. 3, 1323–1335.

Christy, R.J., Kaestner, K.H., Geiman, D.E., Lane, M.D., 1991. CCAAT/enhancer binding gene promoter: Binding of nuclear factors during differentation of 3T3-L1 preadipocytes. Proc. Natl. Acad. Sci. USA 88, 2593–2597.

Church, G.M., Gilbert, W., 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA 81, 1991–1995.

Comb, M., Goodman, H.M., 1990. CpG methylation inhibits proenkaphalin gene expression and binding of the transcription factor AP-2. Nucl. Acids Res. 18, 3975–3982.

Coulondre, C., Miller, J.H., Farabaugh, P.J., Gilbert, W., 1978. Molecular basis of base substitution hotspots in *Escherichia coli*. Nature 274, 775–780.

Darlington, G.J., Wang, N., Hanson, R.W., 1995. C/EBPα: a critical regulator of genes governing integrative metabolic processes. Curr. Opin. Gen. Dev. 5, 565–570.

Descombes, P., Chojkier, M., Lichtsteiner, S., Falvey, E., Schibler, U., 1990. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. Genes Dev. 4, 1541–1551.

Descombes, P., Schibler, U., 1991. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. Cell 67, 569–579.

Diehl, A.M., Yang, S.Q., 1994. Regenerative changes in the C/EBPα

and C/EBPβ expression modulate binding to the C/EBP site in the c-*fos* promoter. Hepatology 19, 447–456.

Eden, S., Cedar, H., 1994. Role of DNA methylation in the regulation of transcription. Curr. Opin. Genet. Dev. 4, 255–259.

Flodby, P., Liao, D.-Z., Blanck, A., Xanthopoulos, G.H., HŠllstršm, I.P., 1995. Expression of the liver-enriched transcription factors C/EBPα, HNF-1, and HNF-4 in preneoplastic nodules and hepatocellular carcinoma in rat liver. Molec. Carcinogen. 12, 103–109.

Friedman, A.D., Landschulz, W.H., McKnight, S.L., 1989. CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. Genes Dev. 3, 1314–1322.

Gardiner-Garden, M., Frommer, M., 1987. CpG island in vertebrate genomes. J. Mol. Biol. 196, 261–282.

Graves, B.J., Johnson, P.F., McKnight, S.L., 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell 44, 565–576.

Grünert, S., Jackson, R.J., 1994. The immediate downstream codon strongly influences the efficiency of utilization of eukaryotic translation initiation codons. EMBO J. 13, 3618–3630.

Hendricks-Taylor, L.R., Darlington, G.J., 1995. Inhibition of cell proliferation by C/EBPα occurs in many cell types, does not require the presence of p53 or Rb, and is not affected by large T-antigen. Nucl. Acids Res. 23, 4726–4733.

Johnson, P.F., Landschulz, W.H., Graves, B.J., McKnight, S.L., 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes Dev. 1, 133–146.

Jones, P.A., Wolkowicz, M.J., Rideout III, W.M., Gonzales, F.A., Marziasz, C.M., Coetzee, G., Tapscott, S.J., 1990. De novo methylation of the MyoD1 CpG island during the establishment of immortal cell lines. Proc. Natl. Acad. Sci. USA 87, 6117–6121.

Kaestner, K., Christy, R., Lane, M.D., 1990. Mouse insulin-responsive glucose transporter gene: Characterisation of the gene and transactivation by the CCAAT/enhancer binding protein. Proc. Natl. Acad. Sci. USA 87, 251–255.

Kagayama, R., Sasai, Y., Nakanishi, S., 1991. Molecular characterization of transcription factors that bind to the cAMP responsive region of the substance P precursor gene. J. Biol. Chem. 266, 15525–15531.

Katz, S., Kowenz-Leutz, E., Müller, C., Meese, K., Ness, S.A., Leutz, A., 1993. The NF-M transcription factor is related to C/EBPβ and plays a role in signal transduction, differentiation and leukemogenesis of avian myelomonocytic cells. EMBO J. 12, 1321–1332.

Kinoshita, S., Akira, S., Kishimoto, T., 1992. A member of the C/EBP family, NF-IL6 beta, forms a heterodimer and transcriptionally synergizes with NF-IL6. Proc. Natl. Acad. Sci. USA 89, 1473–1476.

Kozak, M., 1989. The scanning model for translation: an update. J. Cell Biol. 108, 229–241.

Landschulz, W.H., Johnson, P.F., Adashi, E.Y., Graves, B.J., McKnight, S.L., 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. 2, 786–800.

Legraverend, C., Antonson, P., Flodby, P., Xanthopoulos, K.G., 1993. High level activity of the mouse CCAAT/enhancer binding protein (C/EBPα) gene promoter involves autoregulation and several ubiquitous transcription factors. Nucl. Acids Res. 21, 1735–1742.

Li, E., Bestor, T.H., Jaenisch, R., 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69, 915–926.

Lin, F.-T., MacDougald, O.A., Diehl, A.M., Lane, M.D., 1993. A 30-kDa alternative translation product of the CCAAT/enhancer binding protein α message: Transcriptional activator lacking antimitotic activity. Proc. Natl. Acad. Sci. USA 90, 9606–9610.

Lincoln, A.J., Williams, S.C., Johnson, P.F., 1994. A revised sequence of the rat C/EBP gene. Genes Dev. 8, 1131–1132.

McKnight, S.L., Lane, M.D., Gluecksohn-Waelsch, S., 1989. Is CCAAT/enhancer-binding protein a central regulator of energy metabilism. Genes Dev. 3, 2021–2024.

McKnight, S.L., 1992. CCAAT/enhancer binding protein. In: McKnight, S.L., Yamamoto, K.R. (Eds.) Transcriptional Regulation, ch. 29, pp. 771–795. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Nerlov, C., Ziff, E.B., 1994. Three levels of functional interaction determine the activity of CCAAT/enhancer binding protein-α on the serum albumin promoter. Genes Dev. 8, 350–362.

Nerlov, C., Ziff, E.B., 1995. CCAAT/enhancer binding protein-α amino acid motifs with dual TBP and TFIIB binding ability co-operate to activate transcription in both yeast and mammalian cells. EMBO J. 14, 4318–4328.

Ness, S.A., Kowentz-Leutz, E., Casini, T., Graf, T., Leutz, A., 1993. Myb and NF-M: combinatorial activators of myeloid genes in heterelogous cell types. Genes. Dev. 7, 749–759.

Ossipow, V., Descombes, P., Schibler, U., 1993. CCAAT/enhancer-binding protein mRNA is translated into multiple proteins with different transcription activation potentials. Proc. Natl. Acad. Sci. USA 90, 8219–8223.

Pei, D., Shih, C., 1991. An 'attenuator domain' is sandwiched by two distinct transactivation domains in the transcription factor C/EBP. Mol. Cell. Biol. 11, 1480–1487.

Poli, V., Mancini, F.P., Cortese, R., 1990. IL-6DBP, a nuclear protein involved in interleukin-6 signal transduction, defines a new family of leucine zipper proteins related to C/EBP. Cell 63, 643–653.

Rana, B., Xie, Y., Mischoulon, D., Bucher, N.L.R., Farmer, S.R., 1995. The DNA binding activity of C/EBP transcription factors is regulated in the G_1 phase of the hepatocyte cell cycle. J. Biol. Chem. 270, 18123–18132.

Razin, A., Cedar, H., 1991. DNA methylation and gene expression. Microbiol. Rev. 55, 451–458.

Roman, C., Suso Platero, J., Shuman, J., Calame, K., 1990. Ig/EBP-1: a ubiquitously expressed immunoglobulin enhancer binding protein that is similar to C/EBP and heterodimerizes with C/EBP. Genes Dev. 4, 1404–1415.

Ron, D., Habener, J.F., 1992. CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene transcription. Genes Dev. 6, 439–453.

Ryden, T.A., Beemon, K., 1989. Avian retroviral long terminal repeats bind CCAAT/enhancer-binding protein. Mol. Cell. Biol. 9, 1155–1164.

Sambrook, L., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, Plainview, NY.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74, 5463–5467.

Studier, W.F., Rosenberg, A.H., Dunn, J.J., Dubendorff, J.W., 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Meths Enzymol. 185, 60–89.

Thomassin, H., Hamel, D., Bernier, D., Guertin, M., Belanger, L., Molecular cloning of two C/EBP-related proteins that bind to the promoter and the enhancer of the α_1-fetoprotein gene. Further analysis of C/EBPβ and C/EBPγ. 1992. Nucl. Acids Res. 20, 3091–3098.

Timchenko, N., Wilson, D.R., Taylor, L.R., Abdelsayed, S., Wilde, M., Sawadogo, M., Darlington, G.J., 1995. Autoregulation of the human C/EBPα gene by stimulation of upstream stimulatory factor binding. Mol. Cell. Biol. 15, 1192–1202.

Umek, R.M., Friedman, A.D., McKnight, S.L., 1991. CCAAT-enhancer binding protein: a component of a differentiation switch. Science 251, 288–292.

Vinson, C.R., Sigler, P.B., McKnight, S.L., 1989. Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science 246, 911–916.

Wang, N.D., Finegold, M.J., Bradley, A., Ou, C.N., Abdelsayed, S.V., Wilde, M.D., Taylor, L.R., Wilson, D.R., Darlington, G.J., 1995. C/EBPα is critical for the regulation of carbohydrate and lipid storage. Science 269, 1108–1112.

Wijnholds, J., Muller, E., AB, G., 1991. Oestrogen facilitates the bind-

ing of ubiquitous and liver-enriched nuclear proteins to the apoVLDL II promoter in vivo. Nucl. Acids Res. 19, 33–41.

Williams, S.C., Cantwell, C.A., Johnson, P.F., 1991. A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers in vitro. Genes Dev. 5, 1553–1567.

Xanthopoulos, K.G., Mirkovitch, J., Decker, T., Kuo, C.F., Darnell Jr, J.E., 1989. Cell-specific transcriptional control of the mouse DNA-binding protein mC/EBP. Proc. Natl. Acad. Sci. USA 86, 4117–4121.

Xu, Q., Tata, J.R., 1992. Characterization and developmental expression of *Xenopus* C/EBP gene. Mech. Dev. 38, 69–81.