Nucleic Acids Research, Vol. 20, No. 12 3091–3098

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

Hélène Thomassin+, Diane Hamel, Daniel Bernier, Michel Guertin§ and Luc Belanger*
Le Centre de Recherche en Cancérologie de l'Université Laval, L'Hôtel-Dieu de Québec, Québec G1R 2J6, Canada

Received March 6, 1992; Revised and Accepted May 7, 1992 EMBL accession no. X64403

ABSTRACT

In an attempt to identify proteins that may regulate α_1-fetoprotein (AFP) gene expression, we screened a cDNA expression library from neonatal rat liver with two essential cis-elements of the AFP promoter and enhancer. We isolated two cDNAs which were found to correspond to leucine zipper proteins of the CC-AAT/enhancer binding protein (C/EBP) family: C/EBPβ and C/EBPγ. The three related proteins C/EBP α, β and γ bind with indistinguishable specificity to multiple DNA sites in the promoter and the enhancer of the AFP gene. In addition, C/EBPβ and C/EBPγ readily heterodimerize with each other as well as with C/EBPα. The mRNAs coding for C/EBPβ and C/EBPγ are expressed in a wider variety of rat tissues than C/EBPα mRNA, including yolk sac and fetal liver. The steady-state levels of C/EBP α, β and γ mRNAs increase during liver development, in parallel with their respective gene transcriptional rates. The high levels of C/EBP β and γ mRNAs in rat yolk sac and fetal liver, where C/EBPα is poorly expressed, suggest that C/EBP β and/or γ could be preponderant or early regulators of the AFP gene in these tissues.

INTRODUCTION

During development of higher eukaryotes, the tissue specific and temporal expression of many genes is determined at the level of transcription initiation. Various combinations of transcription factors control the frequency of transcription initiation by binding to specific DNA recognition sites within promoter and enhancer regions of the target genes. Transcription factors implicated in the regulation of tissue-specific gene expression are themselves submitted to different modes of regulation and are often expressed in a cell-type-specific manner (1–3).

The liver is a rich model to study tissue-specific gene expression. Specific expression of several genes in hepatocytes is modulated by various induction and repression mechanisms according to the stage of development, hormonal signals, or pathological conditions (4). For instance, the α_1-fetoprotein (AFP) gene, subject to tight transcriptional controls, is actively expressed in hepatocytes of the fetal and newborn rat liver, and in endodermal cells of the yolk sac. AFP gene expression in liver is gradually shut off around birth to reach almost undetectable levels during normal adult life. Re-expression of the gene can be observed during hepatocarcinogenesis and chemical or physical injuries to the liver (5–7). Moreover, AFP gene transcription can be repressed by glucocorticoid hormones in the developing liver (5, 8).

Studies using transfection into tissue culture cells or transgenic mice have led to the identification and characterization of a proximal promoter and distal enhancers within 7 kb upstream of the transcription start site of the AFP gene (9–16). The region included in the first 250 bp contains information required for both basal and tissue-specific AFP promoter activity (15–17). In the rat AFP promoter, deletions and point mutations have defined three essential elements stringently conserved in human and mouse, PE1, PE2, and DE at −65, −120, and −160 bp, respectively (15). PE1 is recognized by transcription factor Hepatocyte Nuclear Factor 1 (HNF1) which appears to interact with many liver specific promoters (18–21). Element DE contains a glucocorticoid receptor binding site (15), and is thought to mediate, at least in part, the AFP gene response to glucocorticoids (15, 16, 22, 23). Element PE2 is part of a regulatory region which contains overlapping DNA binding sites for transcription factors Nuclear Factor 1 (NF1), CC-AAT/enhancer binding protein (C/EBPα) and HNF1 (17, 18, 24, 25, and our own results) (Figure 1). The rat AFP gene region located between −2 and −4 kb behaves as a typical enhancer, essential and sufficient for high AFP promoter activity of AFP-chloramphenicol acetyltransferase (CAT) constructs transfected into rat or human hepatoma cells (15). More recently, we used a series of AFP-CAT constructs, containing internal deletions within this enhancer region, to locate a 0.7 kb segment (−2.8 to −2.1 kb) sufficient for full-level expression of the reporter gene in hepatoma cell. This DNA segment contains at least 6

* To whom correspondence should be addressed

Present addresses: +Centre de Recherche sur l'Endocrinologie Moléculaire et le Développement, CNRS, UPR 1511, 9 rue Jules Hetzel, 92190 Meudon, France and §Département de Biochimie, Faculté des Sciences et Génie, Université Laval, Québec, Canada

distinct liver nuclear protein binding sites. Three of these sites, EE1, EE8, and EE7 (Figure 1), were found to bind purified C/EBPα and a similar set of unidentified liver nuclear proteins. This binding is competed out by promoter element PE2. Furthermore, C/EBPα was able to transactivate the AFP gene when transfected into hepatoma cells (Bernier, Thomassin, Bélanger et al., unpublished results).

C/EBPα is a transcription factor expressed primarily in liver, fat and intestinal tissues (27). It belongs to a class of proteins, termed bZIP proteins, characterized by a bipartite DNA-binding domain containing a basic amino acids-enriched region that contacts the DNA and an adjacent region called the leucine zipper required for dimerization (28–30). However, the expression of C/EBPα is linked to terminally differentiated cells (27, 31) and is hardly detectable in fetal liver and in yolk sac endodermal cells (32, Figure 7). This argued against a role of C/EBPα in early activation of the AFP gene which is one of the first genes expressed in liver development.

Here we report the cloning of two cDNAs encoding proteins which recognize the AFP sites EE1, EE7, EE8 and PE2. These two proteins are distinct but related to C/EBPα and can heterodimerize with one another as well as with C/EBPα. The expression profiles of their mRNAs suggest that these two factors are earlier transcription regulators than C/EBPα during rat development.

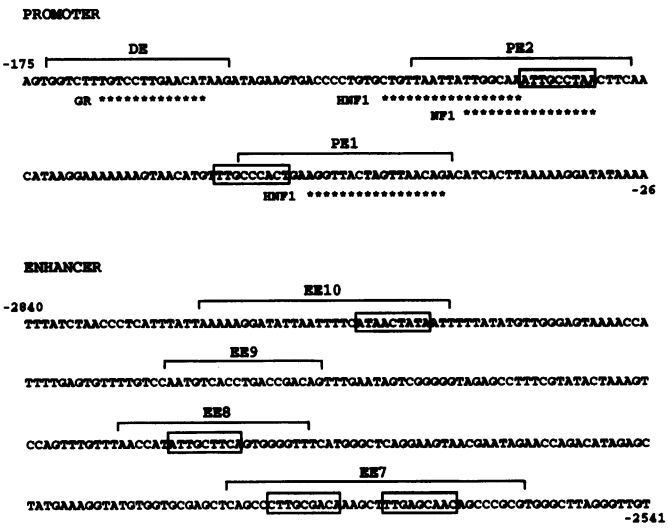

Figure 1. Sequence of rat AFP gene promoter and enhancer domains required for high AFP gene expression. Brackets identify protected regions in DNase1 footprinting experiments with liver nuclear extracts, including four sites (DE, PE2, PE1, EE7) defined as cis-essential by site-specific mutation (Figure 4; 15 and unpublished results). Boxed sequences indicate homology regions with the consensus binding site of C/EBP 5′-TT/GNNGC/TAAT/G-3′ (26). Recognition sequences for the glucocorticoid receptor (GR), NF1 and HNF1 are also identified.

MATERIALS AND METHODS

Isolation of C/EBPβ and C/EBPγ cDNA clones

Liver polyA+ RNA was isolated from 4-day old Sprague-Dawley rats. Ten micrograms of polyA+ RNA were used to synthesize oligo (dT)-primed double stranded cDNA with the reagents and protocols of Pharmacia. EcoR1/Not1 adaptors were added and size-selected cDNA (>500 pb) was cloned into λgt11 bacteriophage. The concatenated DNA was packaged *in vitro* using a commercial extract (Stratagene). A cDNA library of 1.4×10^7 plaques was obtained and was not amplified before use. A total of 7×10^5 recombinant plaques were screened according to Singh et al. (33) and Vinson et al. (34) with ^{32}P-end labeled concatenated oligonucleotides of enhancer site EE7:

5′-gatccAGCTCAGCCCTTGCGACAAAGCTTTGAGCAACA-GCCCGCGTGGGCg-3′
3′-gTCGAGTCGGGAACGCTGTTTCGAAACTCGTTCTCGGGCGCA-CCCGcctag-5′

and of promoter site PE2:

5′-aattcAATTATTGGCAAATTGCCTAACTTCG-3′
3′-gTTAATAACCGTTTAACGGATTGAAGCttaa-5′

Three recombinant phages (L6, L7 and L27) that hybridized with both EE7 and PE2 probes were isolated. Y1089 lysogens harboring L6, L7 or L27 were isolated and were induced to express their respective β-galactosidase fusion protein (35). Y1089 lysogen harboring L20, a C/EBPα recombinant phage (36) kindly provided by S. L. McKnight, was processed similarly. Lysogen extracts were prepared according to (33) and were tested for the presence of specific DNA-binding proteins in a gel retardation assay. DNA sequence analysis indicated that the three cDNA inserts obtained were encoding two different proteins, namely C/EBPβ and C/EBPγ. Full length C/EBPβ cDNA and longer C/EBPγ cDNA clones were isolated by rescreening the same library with the original inserts as hybridization probes according to Sambrook et al. (37). cDNA inserts were subcloned into pBluescript vector (Stratagene) and sequenced on both strands by the dideoxy method (38).

Overexpression and purification of recombinant proteins

An overexpression system of phage T7 promoter-containing vectors (39, Novagen) was used to generate a series of expression plasmids encoding recombinant C/EBPα, C/EBPβ or C/EBPγ. An expression plasmid encoding a C/EBPβ protein lacking only the amino-terminal 21 amino acids was constructed as described by Descombes et al. (40), except for the use of vector pET 11d instead of pET 8c. An expression vector encoding truncated C/EBPγ (carboxy-terminal 211 amino acids) was obtained by inserting a blunted Not1 fragment of L6 cDNA insert into the blunted BamH1 site of phage T7 expression vector pET 11a. A C/EBPα expression plasmid encoding the carboxyl-terminal 143 amino acids of the protein was constructed with a 815 bp Pst1 fragment isolated from expression vector pMSV-EBP (41) and subcloned in the Pst1 site of pBluescript polylinker. The C/EBPα insert was excised from pBluescript polylinker by Not1/EcoR1 digestion, blunted and inserted into the blunted BamH1 site of pET 11c vector. The resulting chimeric plasmids were used to transform the bacterial strand BL 21/DE-3/pLysS. Recombinant proteins were purified from bacterial extracts by fractionation on heparin sepharose columns according to (40).

DNase protection and gel retardation assays

Double-stranded oligonucleotide probes for gel retardation assays, and DNA probes for DNase1 footprinting assays, were end-labeled by filling in of cohesive ends using ^{32}P-labeled deoxinucleotides and Klenow polymerase, and purified by polyacrylamide gel electrophoresis followed by electroelution.

For gel retardation assays, purified recombinant proteins were incubated 30 min on ice in a 10 μl reaction containing 7%

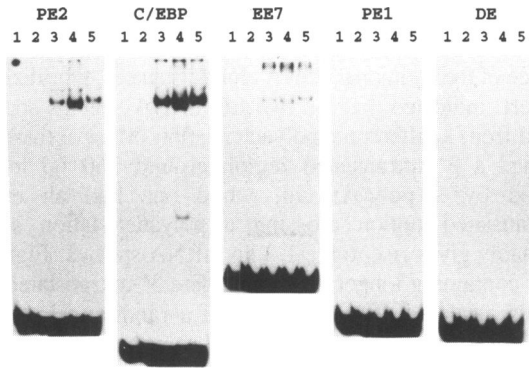

Figure 2. DNA binding assays of proteins encoded by L6 and L7 recombinant phages. Gel retardation assays were carried out with protein extract from Y1089 (lanes 2) and from λ lysogens containing L20 (lanes 3), L6 (lanes 4) or L7 (lanes 5) recombinant phages. Radiolabeled oligonucleotides used as probes were AFP promoter sites PE2, PE1 or DE, and AFP enhancer site EE7, indicated in Figure 1, or a C/EBP synthetic binding site (36). Lanes 1: probes in the absence of protein extract. Similar results were obtained with recombinant phage L27.

glycerol, 17.5 mM Hepes (pH 7.6), 5 mM $MgCl_2$, 80 mM KCl, 0.7 mM EDTA (pH 8.0), 1 mM DTT, 0.035% TritonX100, 1.5 μg of poly(dI-dC) (Pharmacia) and 1 ng of radiolabeled probe. Free DNA and DNA-protein complexes were resolved on a 5 or 7% polyacrylamide gel (acrylamide-bisacrylamide weight ratio 30:1) in 0.25×TBE (25 mM Tris base, 25 mM boric acid and 1 mM EDTA). The gels were pre-electrophoresed at 10 V/cm for 2h at room temperature and the buffer was changed before the samples were loaded. Electrophoresis was performed at the same voltage for 2−3h. Gels were fixed, dried and exposed to X-ray film without intensifying screen.

For DNase1 footprinting assays, purified recombinant proteins (arbitrary units) or up to 40 μg of rat liver nuclear proteins were incubated 30 min on ice with 0.5 to 1 ng of labeled DNA fragment using the same conditions as for gel retardation assays but in a total volume of 40 μl. DNase1 freshly diluted in 2.5 mM $MgCl_2$, 25 mM Tris-HCl (pH 7.7), 10% glycerol and 100 μg/ml BSA, was added in 1−5 μl volume and the mixture was incubated at 25°C for 2 min. Reactions were stopped with 60 μl of 200 mM NaCl, 20 mM EDTA, 1% SDS, 250 μg/ml yeast tRNA and 150 μg/ml proteinase K. Samples were incubated for 60 min at 37°C, extracted with an equal volume of phenol: chloroform, and ethanol precipitated. DNA pellets were dried, resuspended in 80% formamide, heated at 90°C for 3 min and loaded on a 6% polyacrylamide, 8.3 M urea sequencing gel. Gels were directly autoradiographed at −70°C with a Dupont-Cronex intensifying screen.

Nuclear proteins from 4-day old rat liver were extracted according to (42, 43) with the addition of 1% (wt/vol) low-fat dry milk in homogenization buffer as a competitive protease inhibitor (40, 43).

Northern Blot Analysis and Nuclear Run-On Assays

Tissues were removed from 10 to 30 Sprague-Dawley rats and total cellular RNA was extracted using the guanidinium thiocyanate method (44). After selection on oligo d(T) cellulose columns, polyA+ RNA (3 μg) was separated on 1.5% agarose-formaldehyde gel, transfered to a Hybond N membrane (Amersham), and hybridized according to (37). RNA was visualized using ethidium bromide staining as described in (45) to ensure that equivalent amounts of intact RNA were being probed. Partial C/EBPβ and C/EBPγ cDNAs obtained in the original screening, a pMSV-C/EBPα EcoR1-HindIII fragment carrying C/EBPα cDNA, and near full length rat AFP cDNA isolated from pHDQ105 plasmid (46) were labeled by random priming and used as probes. The same relative amounts of C/EBP α, β, γ mRNAs among the tested tissues were reproducibly detected in several blots.

Purification of rat liver nuclei and nuclear run-on assays were performed as described previously (46). Radiolabeled nascent RNA chains were hybridized to an excess of different cDNAs (3 μg dots) immobilized on nitrocellulose filters. Washed and RNase-treated filters were subjected to autoradiography at −70°C for 1−3 days with intensifying screens.

RESULTS

Cloning of two cDNAs encoding proteins that bind to regulatory elements of the AFP gene

As mentioned, we anticipated the binding of other proteins than C/EBPα to the AFP enhancer sites EE1, EE7 and EE8, and to the AFP promoter site PE2. To isolate cDNA clones encoding factors which recognize these sites we directly screened a 4-day old rat liver cDNA expression library in λgt11 with radiolabeled double-stranded EE7 and PE2 oligonucleotides. Both the procedure originally described (33) and its modification (34) were used in parallel since some fusion proteins are suspected to be irreversibly denatured by guanidinium chloride. Screening of 7×10^5 recombinant phage plaques yielded one clone (L6) detected by both procedures (33, 34) and two (L7 and L27) only detected with the modified protocol (34).

Lysogens containing recombinant phages L6, L7 or L27, were isolated along with lysogens for L20 which is a recombinant phage encoding the DNA binding domain of C/EBPα fused to β-galactosidase (36). Partially purified extracts from IPTG-induced lysogens were tested for the presence of specific DNA-binding proteins in gel retardation assays. The β-galactosidase fusion proteins produced by L6, L7, L27 and L20 recombinant phages exhibited the same DNA-binding specificity (Figure 2). Formation of complexes were observed with oligonucleotides containing sites PE2 or EE7, or a binding site for C/EBP (36) but not with oligonucleotides containing the AFP promoter sites PE1 or DE. However, under stringent cross-hybridization conditions, the L6 insert (900 pb) was found to hybridize with L6 but not with L7, L27 and L20 recombinant phages, whereas both L7 and L27 inserts (1000 and 800 bp) strongly hybridized with L7 and L27 phages, slightly with L20 and not with L6 (data not shown). Nucleotide sequence analysis confirmed that the three recombinant phages we had isolated contained derivatives of two mRNA species. The complete sequence of L27 insert was included in the sequence of L7 insert. L6 and L7 cDNAs encoded bZIP DNA-binding proteins distinct from C/EBPα but with strong amino acid sequence homologies to the C/EBPα DNA-binding domain (Figure 3). The sequence of the 858 bp insert of L6 showed a partial open reading frame of 634 bp that contributed 211 amino acids to the β-galactosidase fusion protein. Northern blot analysis of polyA+ liver RNA revealed two mRNA species of 1.3 kb and 3.7 kb that hybridized with the L6 insert (see figures 6 and 7). These two L6 mRNA species derived from a unique single copy gene as indicated by Southern blot analysis (data not shown).

a

```
   1  GGCCGCGGTGGCGGAACGGGCGGAGGCTGCCGGTTTCGTAACCGTCGCTCCTCCTCGCCG   60
      G  R  G  G  G  T  G  G  G  C  R  F  R  N  R  R  S  S  S  P

  61  ACTCGCAGGCTGCGAGGCCTGGGTAGGGTCGGACCGCGCCGCGCGGGGCCGCTCGGAGTG  120
      T  R  R  L  R  G  L  G  R  V  G  P  R  R  A  G  P  L  G  V

 121  GAGGCCGTCTGGGGGCGGGCGGGCCGGCCGGAGGCGCAGGTACATGTGAAGATTTTTTGG  180
      E  A  V  W  G  R  A  G  R  P  E  A  Q  V  H  V  K  I  F  W

 181  CAGTTGAGTGTGGCCTTCTCGGATCACACTGCTCTGATTTCTACGTGTTCTGTGTTGGTG  240
      Q  L  S  V  A  F  S  D  H  T  A  L  I  S  T  C  S  V  L  V

 241  GAGGAACGGGCCCAAATGAGCAAGCTGTCGCAGCCAGCCTCGACTGCAGGAGTGAACGGG  300
      E  E  R  A  Q  M  S  K  L  S  Q  P  A  S  T  A  G  V  N  G

 301  ATCAGTGTCATTCACACTCAGGCACATGCCAGCGGCTTACAGCAGGTTCCTCAGCTGGTG  360
      I  S  V  I  H  T  Q  A  H  A  S  G  L  Q  Q  V  P  Q  L  V

 361  CCCGCTGGGCCTGGGGGAGGAGGCAAAGCTGTGCCTCCAAGCAAGCAAAGCAAGAAGAGC  420
      P  A  G  P  G  G  G  G  K  A  V  P  P  S  K  Q  S  K  K  S

 421  TCACCCATGGATCGAAATAGCGATGAGTACCGCCAGCGCAGAGAGCGGAACAATATGGCT  480
      S  P  M  D  R  N  S  D  E  Y  R  Q  R  R  E  R  N  N  M  A

 481  GTGAAAAAGAGCCGGTTAAAAAGCAAGCAGAAAGCTCAAGATACACTGCAGAGAGTAAAC  540
      V  K  K  S  R  L  K  S  K  Q  K  A  Q  D  T  L  Q  R  V  N

 541  CAGCTCAAGGAGGAGAACGAGCGGCTGGAAGCCAAAATTAAGTTGCTGACAAAGGAATTA  600
      Q  L  K  E  E  N  E  R  L  E  A  K  I  K  L  L  T  K  E  L

 601  AGTGTACTGAAAGACTTGTTTCTTGAGCATGCGCACAACCTCGCAGACAATGTGCAGCCC  660
      S  V  L  K  D  L  F  L  E  H  A  H  N  L  A  D  N  V  Q  P

 661  ATCAGCACGGAAACTACGGTGGCAAATTCTGATAACACAGGGCAGTAGATCTCCTTCCAG  720
      I  S  T  E  T  T  V  A  N  S  D  N  T  G  Q  * 235

 721  GCGCCACAGCCTTGTGACTTGAACATGAGAGGTGTGACCACCCTGCACCCCTTGGTTCAG  780

 781  TGGCTGAGCTCGGTCTTGTTCCATTGGAGGTTATTTTTTAGGCTCAGCACTGGACAGCTG  840

 841  GTTAGCCATGTAATTTATCTGGTCTTAAATGATAATGGATTTTTGAAGCACTAAAAGAAA  900

 901  ACTTAGGGTTTATGGTTAAAATAAAATTCCTGGCCTTTAATATTCTTAATAAATCCTCAC  960

 961  TTCCCCAGAAATGCTTCTTTTGTAGGAAGGGGAAAATGGAAGGTCATATTGAAGGTTCTG 1020

1021  GGTTCTAATGAGATAATACCTTGGACTAGGAAAAGCTGAATTCCATCTGTGTTCTTCAAT 1080

1081  TTGTGATTTTATTCTAAATTGTGTATGATAAATTCTAGGGCTATAGAAACCAGCCATAAT 1140

1141  AATTTGGCCCAATTTTATAGATGTATAAACCTGTGTTAGCATAATCAGCATCGTTTTTTC 1200

1201  CAACACATACACATTTTAAAAAATGACCTTAGGTGAAGTAAAGAAATCCCATCAGGTCAC 1260

1261  GTCGAGAGGTGGTTAAAAATACAAGTTGTGTTTAGGAAGGAAAGAGGTGGTTTTGTGCTT 1320

1321  TTCCCCCTTGTTTGTTTTAAAGGGCATAGAGTGTGATTGAACTTTGGGGGGCCTTCTGCT 1380

1381  TATTGGATGAGCATTATCTGTATCCTAAAAGACAGTAGCCACCGTGTACT 1430
```

b

```
C/EBP β   ATPSGSSGSLSTSSSSSPPGTPSPADAKAAPAACF...AGPPAAPAKAKAK
          :**  *: : :: :::: **: :: :: *:* : :   :*  :::: :***
C/EBP α   PTPVPSPHPAPAMGAAGLPGPGGSLKGLAGPHPDLRTGGGGGGGAGAGKAK
             :: * ::::::: ::  : :  :**: : *:* ::* **** ::  :*
C/EBP γ   MSKLSQPASTAGVNGISVIHTQAHASGLQ.QVPQLVPAGPGGGGKAVPPSK

C/EBP β   KA.....VDKLSDEYKMRRERNNIAVRKSRDKAKMRNLETQHKVLELTAEN
          *:      *** *:**::**************** **:***:******::*
C/EBP α   KS.....VDKNSNEYRVRRERNNIAVRKSRDKAKQRNVETQQKVLELTSDN
          :*     :*:**:*** ******:**:*** *:**:: :* *:* :*:::*
C/EBP γ   QSKKSSPMDRNSDEYRQRRERNNMAVKKSRLKSKQKAQDTLQRVNQLKEEN
                            BB-BN--AA-B-R-BB

C/EBP β   ERLQKKVEQLSRELSTLRNLFKQLPE.PLLASAGHC
          :**:*:******:***::*:*** :*: : *:*
C/EBP α   DRLRKRVEQLSRELDTLRGIFRQLPESSLVKAMGNCA
          :**  ::: *::**::*:::*
C/EBP γ   ERLEAKIKLLTKELSVLKDLFLEHAHNLADNVQPISTETTVANSDNTGQ
```

Figure 3. a. Nucleotide sequence and deduced amino acid sequence of C/EBPγ. The sequence of the longer cDNA insert isolated from the newborn rat liver cDNA library is presented. The open reading frame spans 235 amino acids. The sequence of the L6 original insert starts at nucleotide 72 and ends at nucleotide 930. The putative polyadenylation signal giving rise to the 1.3 kb mRNA species is underlined. b. Comparison of the carboxy-terminal amino acid sequences of C/EBP α, β and γ spanning the basic domain and the leucine zipper region. Amino acids sequences are displayed such that C/EBPβ (top) and C/EBPγ (bottom) are compared with C/EBPα (middle). Asterisks between the sequences indicate identical amino acids and colons, conservative substitutions. The 17 amino acids segment strikingly conserved in the basic DNA binding domain of C/EBP-related proteins except C/EBPγ, is boxed. The leucines organized in heptad repeat and the i+4 positions relative to this repeat are indicated.

To obtain full-length cDNAs, the original inserts were used as probes to rescreen the newborn rat liver cDNA library. The sequence of the additional cDNA clones isolated, hybridizing with L6 insert, indicated that the two L6 mRNA species apparently resulted from a differential polyadenylation. Most of these clones contained a 3'-untranslated region around 260 bp in length followed by a poly(A) tail, while one had an extended 3'-untranslated region crossing a polyadenylation site that presumably gives rise to the 1.3 kb mRNA species (Figure 3a). Clones containing longer but incomplete 3'-untranslated region were also isolated after screening of a random-primed λZAP cDNA library from adult rat liver and were partially sequenced. The nucleotide and predicted amino acid sequences of the L6 clones we isolated from the newborn rat liver cDNA library are shown in Figure 3a. L6 cDNA encodes a polypeptide wich differs by only a few amino-acids from the mouse protein Ig/EBP-1 (47). Comparison of the nucleotide sequence of L6 with that of Ig/EBP-1 reveals the presence of an additional G at position 153 giving a frame shift in the amino-terminal region.

A further recombinant phage isolated from the newborn rat liver cDNA library, and that hybridized with L7, contained an insert of 1482 bp and showed a first ATG codon at nucleotide 45 which formed the start of an open reading frame of 891 bp encoding a full length protein of 297 amino acids (data not shown). The sequence of this insert was identical to the cDNA sequence of a C/EBP-related transcription factor isolated from rat liver during the course of these studies, and named LAP (40) orIL-6DBP (48). This transcription factor has been the object of several other studies in different species or tissues under the name of NF-IL6 (49), AGP/EBP (50) and more recently C/EBPβ (51) and CRP2 (52).

Following a nomenclature recently proposed for C/EBP-related proteins (51), proteins encoded by L7 and L6 are refered to as C/EBPβ and C/EBPγ, respectively. A comparison of the deduced amino acid sequences of the three related proteins C/EBP α, β and γ is presented in Figure 3b. The carboxy-terminal regions of the three proteins display a strong homology, more pronounced in the basic region contacting the DNA than in the leucine zipper region required for dimerization. From the number of positive plaques recovered from 2×10^6 recombinants in the newborn liver cDNA library, we deduced a relative frequency of 9×10^{-6} for C/EBPβ mRNA and 6×10^{-6} for the 1.3 kb C/EBPγ mRNA, in the 4-day old rat liver.

C/EBPα, C/EBPβ and C/EBPγ bind to the same elements within the promoter and the enhancer of the AFP gene

Comparison of the basic DNA-binding regions revealed an extensive sequence similarity between the three C/EBP-related proteins. However, C/EBPγ shares only 13 of 17 identical amino acids in the segment containing the consensus bZip basic motif (30, 53). This 17 amino acids segment is strikingly conserved in C/EBPα and C/EBPβ and in all the additional members of the family described recently (51, 52). This difference between C/EBPγ and both C/EBPα and C/EBPβ led us to examine if the three C/EBP-related proteins shared DNA-binding specificities in the promoter and the enhancer of the AFP gene using DNase 1 protection assays. An overexpression system under the control of a phage T7 promoter was used to prepare large amounts of purified recombinant C/EBP-related proteins (see Materials and Methods). The three recombinant peptides occupied the same multiple sites with similar relative affinities among sites

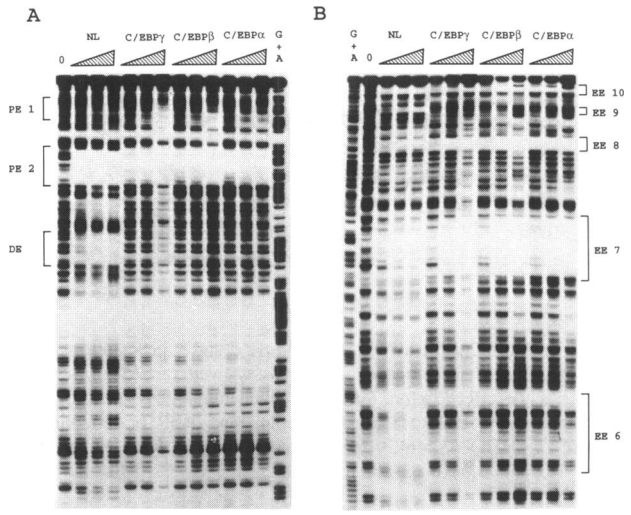

Figure 4. DNase 1 protected regions obtained with newborn liver nuclear proteins and C/EBP-related recombinant proteins, in the rat AFP promoter (A) and enhancer (B) domains. DNase 1 footprinting experiments were conducted with comparable arbitrary footprinting units (1, 5 and 25) of C/EBP γ, β and α and with unfractionated newborn rat liver nuclear proteins (NL: 10, 20 and 40 μg). The HindIII-HaeII promoter fragment (−244 to −15) was isolated from plasmid p1 (15) and end-labeled on the non-coding strand at the HindIII site. The Spe1-BstE2 fragment encompassing 400 bp of the AFP enhancer (−2845 to −2445) was end-labeled on the coding strand at the BstE2 site. Probes submitted to G+A chemical cleavage were used as markers. 0, probes incubated in the absence of test proteins. Brackets indicate the promoter sites PE1, PE2, DE and the enhancer sites EE6 to EE10.

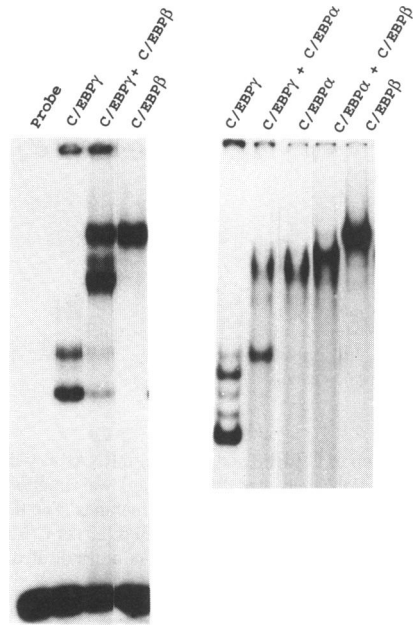

Figure 5. Detection of heterodimers formation between C/EBP α, β and γ. Purified recombinant polypeptides corresponding to C/EBP α, β or γ alone or in the indicated combinations, were incubated with a radiolabeled double-stranded oligonucleotide containing the AFP promoter site PE2. Protein-DNA complexes were resolved by electrophoresis in a non-denaturing polyacrylamide gel.

(Figure 4). In the promoter, besides PE2, two lower affinity sites were detected at −75 bp, overlapping PE1, and at −205 bp. In the enhancer, besides EE7, two additional sites (EE8 and EE1) had high affinity for the three C/EBP peptides and EE10 had an intermediate affinity. Further sites appeared at high protein/DNA ratios but were not detected with rat liver nuclear protein extracts. Also, the protection patterns formed by liver extracts over some sites differed from the one formed by the recombinant peptides (for example, NF1 in liver extracts overrides C/EBP binding to PE2). Some caution is thus in order with the *in vivo* interpretation, but the results demonstrate that the tested members of the C/EBP family have the capacity to bind with similarly high relative affinities to multiple sites in AFP regulatory domains. It is noteworthy that the DNA protection pattern of C/EBPγ was identical to C/EBP α and β, despite divergences in the consensus bZip basic DNA-binding motif (Figure 3b).

Formation of heterodimers between C/EBPα, C/EBPβ and C/EBPγ

The leucine zipper region of the three C/EBP-related proteins contains a heptad repeat of leucines in perfect register relative to the location of the highly conserved basic motif. The leucine zipper regions of C/EBPα and C/EBPβ share almost complete homology when including conservative substitutions. The leucine zipper sequence of C/EBPγ diverges to some extent from C/EBPα and C/EBPβ with a heptad repeat of 4 leucines instead of 5. However, C/EBPγ shares identical amino acids with both C/EBPα and C/EBPβ at the i+4 positions relative to the heptad repeat except a conservative substitution valine/isoleucine (Figure 3b). The i+4 positions are thought to be aligned on the same helical face with the leucines repeat and to contribute to the hydrophobic dimer interface (28). These data suggested that C/EBP α, β and γ might form heterodimers.

To test this, we used phage T7 promoter-containing plasmids that encoded recombinant polypeptides with a molecular size of 20, 36 and 28 kD corresponding to truncated C/EBPα, C/EBPβ and C/EBPγ respectively. The overexpressed recombinant polypeptides were purified by chromatography on heparin sepharose columns. Formation of mixed dimers between the three C/EBP-related proteins was tested in gel retardation assays using site PE2 as a probe (Figure 5). When tested alone, truncated C/EBPα and C/EBPβ formed a complex of unique mobility with the radiolabeled DNA probe. C/EBPγ was particularly labile and only partially degraded C/EBPγ could be isolated by heparin-sepharose chromatography. C/EBPγ resulted in the formation of several complexes with higher mobility than the 20kD form of C/EBPα. When two of the three C/EBP-related proteins were mixed in any combination, complexes of intermediate mobilities were systematically detected. Homodimerization of C/EBPα is required for its binding to DNA (29) and C/EBPα-C/EBPβ heterodimers can be formed as efficiently as C/EBPα or C/EBPβ homodimers (40, 48). Since element PE2 carries a single C/EBP binding site, intermediate PE2 complexes in the gel retardation assays were interpreted to represent heterodimeric complexes.

Expression patterns of C/EBP α, β and γ mRNAs

The presence of C/EBPβ and C/EBPγ mRNAs was investigated by Northern blot hybridization in tissues from 4-day old rats, the time point selected for the library (Figure 6). C/EBPβ mRNA was detected at varying levels as a single 1.5 kb band in all tissues examined except brain (Figure 6). This tissue distribution pattern

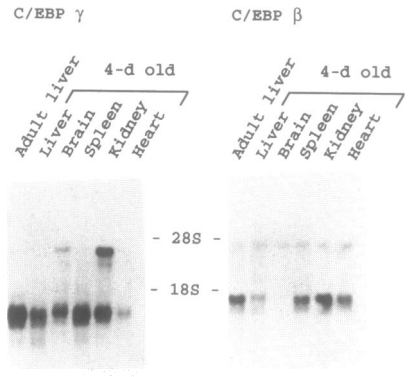

Figure 6. Distribution of C/EBPβ and C/EBPγ mRNAs in 4-day old rat tissues and adult rat liver. Polyadenylated RNA (3 μg) was fractionated on agarose electrophoresis gels and transfered to nylon membranes. The filters were probed under stringent conditions with radiolabeled C/EBPβ (L7) or C/EBPγ (L6) cDNAs. The positions of 28S and 18S ribosomal RNAs are indicated.

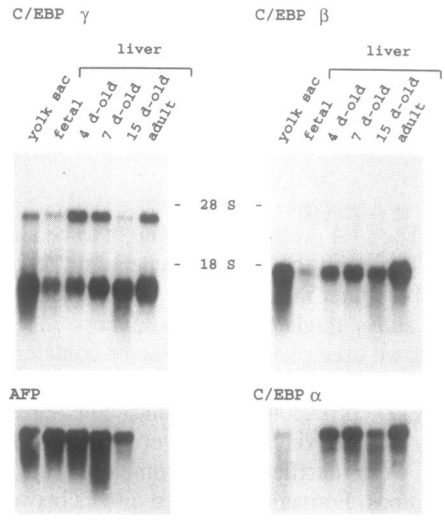

Figure 7. Analysis of C/EBP α, β and γ mRNAs in yolk sac and during liver development. Total RNA was extracted from yolk sac and fetal liver at embryonic day 18 and from 4-, 7-, 15-day old and adult rat liver. PolyA+ RNA was enriched on oligo(dT) cellulose and 3μg aliquots were fractionated on agarose electrophoresis gels, then transfered to nylon membranes and probed with radiolabeled C/EBPγ (L6) or C/EBPβ (L7) cDNA fragments. Filters were dehybridized and reprobed with radioactive C/EBPα and AFP cDNA fragments. The positions of 28 S and 18 S ribosomal RNAs are indicated.

appeared similar to that described in adult rat (40, 48). The 1.3 and 3.7 kb C/EBPγ mRNA species were both found in all tissues examined with a systematic predominance of the short form (Figures 6 and 7). This differs from the situation in the mouse (47, and our own results with mouse liver and L cells) in which the larger form is the major mRNA species detected. This may reflect different uses of polyadenylation signals and/or differences in mRNA stability between rat and mouse.

The accumulation of C/EBP α, β and γ mRNAs was also examined in the yolk sac and as a function of developmental stage in liver (Figure 7). Northern analysis indicated a high abundance of C/EBP β and γ mRNAs in the yolk sac, contrasting with the

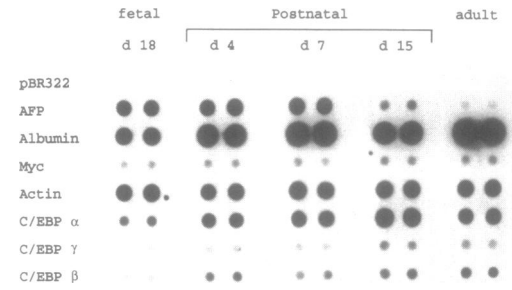

Figure 8. Transcriptional activities of c/ebp α, β and γ genes during rat liver development. RNA chain elongation (run-on) assays were carried out with purified liver nuclei as described before (46) using filter-immobilized pBR322 plasmid as a control, or recombinant plasmids containing AFP or albumin cDNA (pHDQ105 or pHDQ835, 46), v-myc or rat actin cDNA (gifts of N. Marceau) and 1.8, 1.0 and 0.9 kb cDNA inserts of C/EBPα, C/EBPβ and C/EBPγ respectively.

low level of C/EBPα mRNA. C/EBPα mRNA was found only in trace amounts in fetal liver at day 18 of gestation, while C/EBPβ mRNA was easily detected and C/EBPγ mRNA was yet more abundant. All mRNAs increased post-natally, at terminal stages of liver differentiation. The findings on C/EBPα in fetal liver and yolk sac are similar to the situation in the mouse (32). The overall results strengthen the conclusion that C/EBPα has a more restricted pattern of expression than other members of the family, being predominantly linked with mature non growing hepatocytes and adipocytes (27, 31, 51, 54).

To determine whether the observed variations in mRNA accumulation for the three C/EBP-related proteins during development were related to differential transcription of the genes encoding these factors, transcriptional 'run-on' assays were carried out with nuclei isolated from rat liver at different stages. Transcription signals were related to mRNA levels: they increased for c/ebp α, β and γ in terminally differentiating liver, parallel to similar increases in the steady-state level of the corresponding mRNAs. These results thus indicate that increasing levels of C/EBP α, β and γ mRNAs during liver development are transcriptionally controled. Transcription signals were stronger for c/ebpα than for c/ebp β or γ. Since C/EBPα mRNA accumulates almost like C/EBP β and γ mRNA in adult liver (48, 51, and figure 7), and is present in trace amounts in fetal liver, C/EBPα transcript must have a shorter half-life than C/EBP β and γ transcripts.

DISCUSSION

Towards the identification of transcription factors that may regulate the AFP gene, we have isolated the cDNAs for two DNA-binding proteins members of the bZip transcription factors family, C/EBPβ and C/EBPγ, which are closely related to C/EBPα. The three proteins bind with similar specificities to multiple sites within the promoter and the enhancer of the AFP gene and can readily form heterodimers. This is consistent with the amino acids conservation between their carboxy-terminal part. The distribution and developmental pattern of C/EBP α, β and γ mRNAs indicate broader and earlier accumulation of C/EBP β and γ than C/EBPα. Steady-state levels of C/EBP α, β and γ mRNAs are related to transcription rates and increase during liver maturation.

During the course of these studies, several groups independently reported the cloning of cDNAs encoding C/EBP-related proteins. C/EBPγ differs by only a few amino acids from Ig/EBP-1 which was isolated by DNA-ligand mediated screening of a cDNA expression library from mouse fibroblasts using an immunoglobulin heavy-chain gene enhancer element as a probe (47). C/EBPβ has been cloned by direct screening of cDNA expression libraries with regulatory cis-elements from several liver-expressed genes: the D binding site from the serum albumin gene promoter (LAP, 40), IL-1 or IL-6 responsive elements from acute-phase genes (NF-IL6, 49; IL-6DBP, 48) and a cis-acting element from the α1-acid glycoprotein gene promoter (A-GP/EBP, 50). With a similar approach, we have isolated C/EBPβ and C/EBPγ cDNAs from a 4-day old rat liver cDNA expression library using as ligands DNA-binding sites EE7 and PE2 which are important for full activity and tissue-specific function of the enhancer and promoter of the AFP gene (15 and manuscript in preparation).

C/EBPα is predominantly expressed in liver and adipocytes (27, 54) and is thought to play a role in activating target genes specifically expressed in these tissues (41, 55, 56). The isolation of C/EBPβ with DNA ligands from a variety of liver-active promoters, including AFP, readily suggests that this factor, like C/EBPα, is involved in liver-specific gene expression even if C/EBPβ mRNA is expressed at various levels in a wider variety of tissues than C/EBPα (40, 48, 51, 52). Furthermore, the cellular distribution of C/EBPβ appears to be regulated by a post-transcriptional mechanism and the protein preferentially accumulates in liver nuclei (40). In contrast with C/EBP α and β, C/EBPγ transcripts were detected in all tissues examined (Figures 6, 7; 47). Roman et al. (47) reported a preferential accumulation of mouse Ig/EBP-1 mRNA in early B cells, but here we detected C/EBPβ and C/EBPγ mRNAs to a comparable level in rat liver. Thus, although we have no data concerning the accumulation of C/EBPγ protein, the abundance of its cognate mRNA suggests that the three C/EBP-related proteins are present in rat liver.

One approach to the functions of the different C/EBP-related proteins in AFP gene transcription, was to consider their temporal pattern of expression during development. The poor expression of C/EBPα in fetal liver and yolk sac (Figure 7; 32), preponderant sites of AFP synthesis, suggests that C/EBPα is not the major C/EBP-related protein that may regulate the AFP gene, at least in early development. Albumin more than AFP gene expression appears to be correlated with C/EBPα, from the parallel increase in C/EBPα and albumin transcription rates during rat liver development (Figure 8) and in cultured primary hepatocytes (54). The accumulation of C/EBPα at late stages of development supports the notion that this factor is preferentially linked with terminally differentiated, growth-arrested cells (27, 31, 51, 55).

The increase in mRNA concentration and transcription rate for C/EBPβ and C/EBPγ during liver development (Figures 7, 8) indicates that, like C/EBPα, these two proteins might accumulate to maximal level in terminally differentiated hepatocytes. Schibler and co-workers have verified this hypothesis concerning C/EBPβ (57). On the other hand, C/EBPβ and C/EBPγ mRNAs are considerably higher in concentration than C/EBPα mRNA in fetal liver and yolk sac (Figure 7) suggesting that C/EBPβ and C/EBPγ proteins are expressed earlier than C/EBPα during development. Expression of the c/ebpβ gene also takes place very early during adipocytes differentiation, preceding that of c/ebpα (51). Although the expression of C/EBPβ is at least partly regulated at the post-transcriptional level (40), and the detection of C/EBPβ mRNA in yolk sac and fetal liver may not correlate with the level of the corresponding protein in these tissues, Descombes and Schibler (57) did detect C/EBPβ in liver nuclei from fetuses one day before birth. Interestingly, the same group has also provided indications that two proteins are translated from the single 1.5 kb C/EBPβ mRNA: a transcriptional activator and a transcriptional repressor, the ratio of which increases during liver differentiation (57).

It will be difficult to pinpoint the exact implications of the C/EBP-related proteins in AFP gene control, compounded by the increasing number of identified C/EBP-related proteins and the possibilities of heterodimerization between the different family members. If several C/EBP-related proteins are present simultaneously in AFP permissive cells, which protein will occupy a cis-acting regulatory element may depend on the ratio of the different proteins, the stability of the homodimers or heterodimers and their affinity for the specific DNA site. Furthermore, specific cooperative interactions with other transcription factors may be determinant, and it is also possible that, as a consequence of changing concentrations of the different family members during development, a C/EBP-related protein bound to the regulatory regions of the AFP gene might be replaced by another one. At this point, the levels of C/EBP β and γ mRNAs in fetal liver and yolk sac suggest that C/EBP β and γ could be prevailing or early regulators of the rat AFP gene.

ACKNOWLEDGEMENTS

We thank S.L.McKnight for his generous gift of C/EBPα recombinant phage and pMSV-EBP plasmid and N.Marceau for v-myc and actin recombinant DNA. This work was supported by a grant from the Medical Research Council of Canada. H.T. was supported by a fellowship from the Medical Research Council of Canada and D.B. by a studentship from the Fonds de la Recherche en Santé du Québec.

REFERENCES

1. Maniatis, T., Goodbourn, S. and Fischer, J. A. (1987) *Science*, **236**, 1237–1245.
2. Mitchell, P. and Tjian, R. (1989) *Science*, **245**, 371–378.
3. Falvey, E. and Schibler, U. (1991) *FASEB J.*, **5**, 309–314.
4. De Simone, V. and Cortese, R. (1988) InMcLean, N. (ed.), Oxford Surveys on Eukaryotic genes, Oxford University Press, Oxford, Vol. 5, pp 51–90.
5. Bélanger, L., Baril, P., Guertin, M., Gingras, M. C., Gourdeau, H., Anderson, A., Hamel, D. and Boucher, J. M. (1983) *Adv. Enz. Regul.*, **21**, 73–99.
6. Tamaoki, T. and Fausto, N. (1984) In Stein, G. and Stein, J. (ed.) Recombinant DNA and cell proliferation. Academic Press, Inc., New York, pp. 145–168.
7. Tilghman, S.M. (1985) In McLean, N. (ed.), Oxford surveys on eukaryotic genes, Oxford University Press, Oxford, vol. 2, pp. 160–206.
8. Guertin, M., Baril, P., Bartkoviak, J., Anderson, A. and Bélanger, L. (1983) *Biochemistry*, **22**, 4296–4302.
9. Muglia, L. and Rothman-Denes, L. B. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7653–7657.
10. Widen, S. G. and Papaconstantinou, J. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8196–8200.
11. Godbout, R, Ingram, R. and Tilghman, S. M. (1986) *Mol. Cell. Biol.*, **6**, 477–487.
12. Hammer, R. E., Krumlauf, R., Camper, S. A., Brinster, R. L. and Tilghman, S. M. (1987) *Science*, **235**, 53–58.

13. Watanabe, K., Saito, A. and Tamaoki, T. (1987) *J. Biol. Chem.*, **262**, 4812–4818.
14. Godbout, R., Ingram, R. S. and Tilghman, S. M. (1988) *Mol. Cell. Biol.*, **8**, 1169–1178.
15. Guertin, M., Larue, H., Bernier, D., Wrange, O., Chevrette, M., Gingras, M. C. and Bélanger L. (1988) *Mol. Cell. Biol.*, **8**, 1398–1407.
16. Poliard, A., Bakkali, L., Poiret, M., Foiret, D. and Danan, J. L. (1990) *J. Biol. Chem.*, **265**, 2137–2141.
17. Feuerman, M. H., Godbout, R., Ingram, R. S. and Tilghman, S. M. (1989) *Mol. Cell. Biol.*, **9**, 4204–4212.
18. Courtois, G. S., Baumhueter, S. and Crabtree, G. (1988) *Proc. Natl. Acad. Sci. USA* , **85**, 7937–7941.
19. Cereghini, S., Blumenfeld, M. and Yaniv, M. (1988) *Genes Dev.*, **2**, 957–974.
20. Lichtsteiner, S. and Schibler, U. (1989) *Cell*, **57**, 1179–1187.
21. Frain, M., Swart, G., Monaci, P., Nicosia, A., Stämpfli, S., Frank, R. and Cortese, R. (1989) *Cell*, **59**, 145–157.
22. Houart, C., Szpirer, J. and Szpirer, C. (1990) *Nucleic Acids res.*, **18**, 6277–6282.
23. Zhang, X. K., Dong, J. M. and Chiu, J. F. (1991) *J. Biol. Chem.*, **266**, 8248–8254.
24. Jose-Estanyol, M. and Danan, J. L. (1988) *J. Biol. Chem.*, **263**, 10865–10871.
25. Zhang, D. E., Hoy, P. R. and Papaconstantinou, J. (1990) *J. Biol. Chem.*, **265**, 3382–3391.
26. Ryden, T. A. and Beemon, K. (1989) *Mol. Cell. Biol.*, **9**, 1155–1164.
27. Birkenmeier, E. H., Gwynn, B., Howard, Jerry, J., Gordon, J. I., Landschulz, W. H. and McKnight, S. L.(1989) *Genes Dev.*, **3**, 1146–1156.
28. Landschulz, W. H., Johnson, P. F. and McKnight, S. L. (1988) *Science*, **240**, 1759–1764.
29. Landschulz, W. H., Johnson, P. F. and McKnight, S. L. (1989) *Science*, **243**, 1681–1688.
30. Vinson, C. R., Sigler, P. B. and McKnight, S. L. (1989) *Science*, **246**, 911–916.
31. Umek, R. M., Friedman, A. D. and McKnight, S. L. (1991) *Science*, **251**, 288–292.
32. Kuo, C. F., Xanthopoulos, K. G. and Darnell,J. E. Jr. (1990) *Development*, **109**, 473–481.
33. Singh, H., Lebowitz, J. H., Baldwin, A. S. Jr and Sharp, P. A. (1988) *Cell*, **52**, 415–423.
34. Vinson, C. R., LaMarco, K. L., Johnson, P. F., Landschulz, W. H. and McKnight, S. L. (1988) *Genes Dev.*, **2**, 801–806.
35. Huynh, T. V., Young, R. A. and Davis, R. W. (1985) In Glover, D. M. (ed.), DNA Cloning—A Practical Approach. IRL Press, Oxford, Vol. 1, pp. 49–78.
36. Landschulz, W. H., Johnson, P. F., Adashi, E. Y., Graves, B. J. and McKnight, S. L. (1988) *Genes Dev.*, **2**, 786–800.
37. Sambrook, J., Fritsh, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Second edition. Cold Spring Harbor Laboratory Press.
38. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
39. Studier, F. W. and Moffart, B. A. (1986) *J. Mol. Biol.*, **189**, 113–130.
40. Descombes, P., Chojkier, M., Lichtsteiner, S., Falvey, E. and Schibler, U. (1990) *Genes Dev.*, **4**, 1541–1551.
41. Friedman, A. D., Landschulz, W. H. and McKnight, S. L. (1989) *Genes Dev.*, **3**, 1314–1322.
42. Gorski, K., Carneiro, M. and Schibler, U. (1986) *Cell*, **47**, 767–776.
43. Maire, P., Wuarin, J. and Schibler, U. (1989) *Science*, **244**, 343–346.
44. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979) *Biochemistry*, **24**, 5294–5299.
45. Rosen, K. M. and Villa-Komaroff, L. (1990) *Focus*, **12**, 23–24.
46. Turcotte, B., Guertin, M., Chevrette, M. and Bélanger, L. (1985) *Nucleic Acids Res.*, **13**, 2387–2398.
47. Roman, C., Platero, J. S., Shuman, J. and Calame, K. (1990) *Genes Dev.*, **4**, 1404–1415.
48. Poli, V., Mancini, F. P. and Cortese, R. (1990) *Cell*, **63**, 643–653.
49. Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., Nakajima, T., Hirano, T. and T. Kishimoto, T. (1990) *EMBO J.*, **9**, 1897–1906.
50. Chang, C. J., Chen, T. T., Lei, H. Y., Chen, D. S. and Lee, S. C. (1990) *Mol. Cell. Biol.*, **10**, 6642–6653.
51. Cao, Z., Umek, R. M. and McKnight, S. L.(1991) *Genes Dev.*, **5**, 1538–1552.
52. Williams, S. C., Cantwell, C. A. and Johnson, P. F. (1991) *Genes Dev.*, **5**, 1553–1567.
53. Kouzarides, T. and Ziff, E. (1988) *Nature*, **336**, 646–651.
54. Xanthopoulos, K. G., Mirkovitch, J., Decker, T., Kuo, C. F. and Darnell, J. E. Jr. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 4117–4121.
55. Christy, R. J., Yang, V. W., Ntambi, J. M., Geiman, D. E., Landschulz, W. H., Friedman, A. D., Nakabeppu, Y., Kelly, T. J. and Lane, M. D. (1989) *Genes Dev.*, **3**, 1323–1335.
56. Kaestner, K. H., Christy, R. J. and Lane, M. D. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 251–255.
57. Descombes, P. and Schibler, U. (1991) *Cell*, **67**, 569–579.