The EMBO Journal vol.10 no.6 pp.1435-1443, 1991

LFB3, a heterodimer-forming homeoprotein of the LFB1 family, is expressed in specialized epithelia

Vincenzo De Simone, Luisa De Magistris, Domenico Lazzaro, Jutta Gerstner, Paolo Monaci[1], Alfredo Nicosia[1] and Riccardo Cortese[1]

European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, FRG and [1]Istituto di Ricerche di Biologia Molecolare (IRBM) Via Pontina, Km 30.600, 00040 Pomezia, Roma, Italy

Communicated by R.Cortese

We have cloned and characterized a mouse cDNA coding for LFB3, a DNA binding protein containing an extra-large homeodomain. The first 315 amino acids of LFB3 are highly homologous to the DNA binding domain of LFB1, a regulatory protein involved in the expression of several liver-specific genes. LFB3 is a transcriptional activator which binds to DNA as a dimer and forms heterodimers with LFB1 both *in vitro* and *in vivo*. However, LFB3 expression seems not to be directly correlated with the liver-specific phenotype, since it is detected in dedifferentiated hepatoma cell lines which express neither LFB1 nor several liver-specific genes. LFB3 expression starts before that of LFB1 during mouse and rat development, and is strongly increased upon retinoic acid induced differentiation of F9 embryonic carcinoma cells. LFB3 and LFB1 are expressed in the epithelial component of many organs of endodermal and mesodermal origin, suggesting that they may play a more general role associated with the differentiation of specialized epithelia.

Key words: *trans*-acting factors/liver-specificity/homeoproteins/epithelial cells/heterodimers.

Introduction

Hepatocytes are highly specialized epithelial cells responsible for the synthesis of a large number of liver-specific proteins, usually grouped into broad classes according to their role in the general metabolism (Arias *et al.*, 1982). The regulation of liver-specific gene expression is mediated by *cis*-acting elements which serve as binding sites for *trans*-acting regulatory proteins (for a review see De Simone and Cortese, 1988). The genes coding for several transcription factors which play a pivotal role in the activation of many liver-specific genes have been isolated and characterized; these include C/EBP (Landschulz *et al.*, 1988), LFB1/HNF1 (Frain *et al.*, 1989), IL6DBP/LAP (Poli *et al.*, 1990; Descombes *et al.*, 1990), DBP (Mueller *et al.*, 1990) and HNF3 (Lai *et al.*, 1990).

Mutational analysis points to LFB1 as being a crucial transcriptional activator required for the expression of several liver-specific genes *in vitro*, in cultured cell lines and in transgenic mice (Feuerman *et al.*, 1989; Tronche *et al.*, 1989; Courtois *et al.*, 1987 and 1988;; Monaci *et al.*, 1988; Schorpp *et al.*, 1988; De Simone *et al.*, 1987; Lichtsteiner *et al.*, 1987; Tripodi and Cortese, unpublished observations). We and others reported the isolation and characterization of cDNA clones coding for LFB1 (Frain *et al.*, 1989; Baumhueter *et al.*, 1990) and the mapping of its dimerization, DNA binding and transcriptional activation domains (Nicosia *et al.*, 1990). We also showed that LFB1 forms dimers in solution, and that dimerization is required for the binding to DNA. The DNA binding region of LFB1 is characterized by a new type of homeodomain, having 21 extra amino acids compared with the 'canonic' Antennapedia homeodomain (Scott *et al.*, 1989; Finney, 1990; Nicosia *et al.*, 1990).

The presence of LFB1 in rat hepatoma cell lines is directly correlated to the expression of the liver phenotype. Weiss and coworkers described a set of rat hepatoma cell lines which consists of two lines (H4II and C2rev7), which retain a fully differentiated phenotype, and two 'dedifferentiated' subclonal derivatives (H5 and C2) which have lost the capacity to express several liver-specific genes (Deschatrette and Weiss, 1980). The LFB1 gene is only expressed in the fully differentiated H4II and C2rev7 and not in the dedifferentiated H5 and C2 cell lines (Cereghini *et al.*, 1990). However, another DNA binding activity, called vHNF1 (Baumhueter *et al.*, 1988) or vAPF (Cereghini *et al.*, 1988), was identified in the dedifferentiated H5 and C2 rat hepatoma cell lines or in cell hybrids between hepatoma and fibroblast cell lines, which have lost the ability to express some liver-specific genes and, concomitantly, have no LFB1 (Cereghini *et al.*, 1990). Reversion of dedifferentiated cells to the hepatic phenotype was found to be associated with the loss of vHNF1 and the reappearance of LFB1, as if the two activities were mutually exclusive.

In this paper we report that the vHNF1 activity in dedifferentiated hepatoma cells consists of two components, LFBu and LFB3. We have cloned a cDNA coding for LFB3, and we show that it encodes a transcriptional activator which also contains an extra-large homeodomain and forms heterodimers with LFB1 both *in vitro* and *in vivo*. The expression pattern of LFB3 suggests that it may, together with LFB1, play a role in epithelial differentiation.

Results

Identification of a DNA binding protein immunologically related to LFB1 in dedifferentiated hepatoma cell lines

We have investigated the relationship between LFB1 and vHNF1 by band-shift experiments, using nuclear extracts from H5 and C2 cell lines, as well as from their differentiated counterparts H4II and C2rev7. Using high resolution polyacrylamide gel fractionation, we resolved the single band found in H5 cells into two distinct protein–DNA complexes, indicated as LFB3 and LFBu (Figure 1, lane 5). LFB3 is only detected in H5, whereas LFBu is also present in H4II

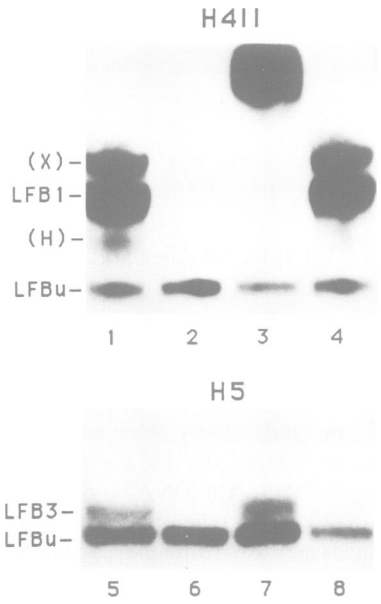

Fig. 1. DNA binding assay and antibody competition with H4II and H5 nuclear extracts. The double stranded oligonucleotide probe contains the LFB1 recognition sequence of the human α1-antitrypsin promoter (Nicosia *et al.*, 1990). Only the retarded bands are shown. 1 and 5 = band-shift in the absence of antibodies. 2 and 6 = + αB1-A. 3 and 7 = + αB1-B. 4 and 8 = + αB3 polyclonal antibodies.

(Figure 1, lane 1), albeit at lower levels. Further investigation indicated that LFBu is present in every nuclear extract tested so far, hence the name LFBu (ubiquitous). Nuclear extracts from the H4II cells display, in addition to LFBu, three band-shifts indicated in Figure 1 as LFB1, X and H. C2 and C2rev7 gave patterns that were identical to H5 and H4II respectively (data not shown).

A polyclonal antibody recognizing the DNA binding domain of LFB1 (αB1-A) specifically competes the LFB3 but not LFBu complex (Figure 1, lane 6). Conversely, another polyclonal antibody (αB1-B) recognizing a C-terminal peptide of the LFB1 activation domain ADI does not compete either LFB3 or LFBu (Figure 1, lane 7). Both αB1-A and αB1-B antibodies abolish the formation of LFB1−X and −H complexes in H4II (Figure 1, lanes 2 and 3); αB1-B addition results in the appearance of a slower migrating band, presumably representing the antibody−protein−DNA complex. In all cases the LFBu complex is never affected by either αB1-A or αB1-B antibodies. These results suggest that LFB3 is homologous to LFB1 in the DNA binding region, but differs in the C-terminal part; LFB1−X and −H complexes must contain LFB1, since they are obliterated by both antibodies.

Isolation of LFB3 cDNA clones reveals significant homologies between LFB3 and LFB1

Since LFB1 mRNA is absent from H5 and C2 cells (Cereghini *et al.*, 1990), we can exclude the possibility that LFB3 is a degradation product of LFB1. On the assumption that the conserved epitopes might correspond to a conserved cDNA sequence, we decided to clone the corresponding cDNA. A set of degenerated oligonucleotides were used for PCR amplification of cDNA segments coding for the LFB1 DNA binding domain and related sequences. The results

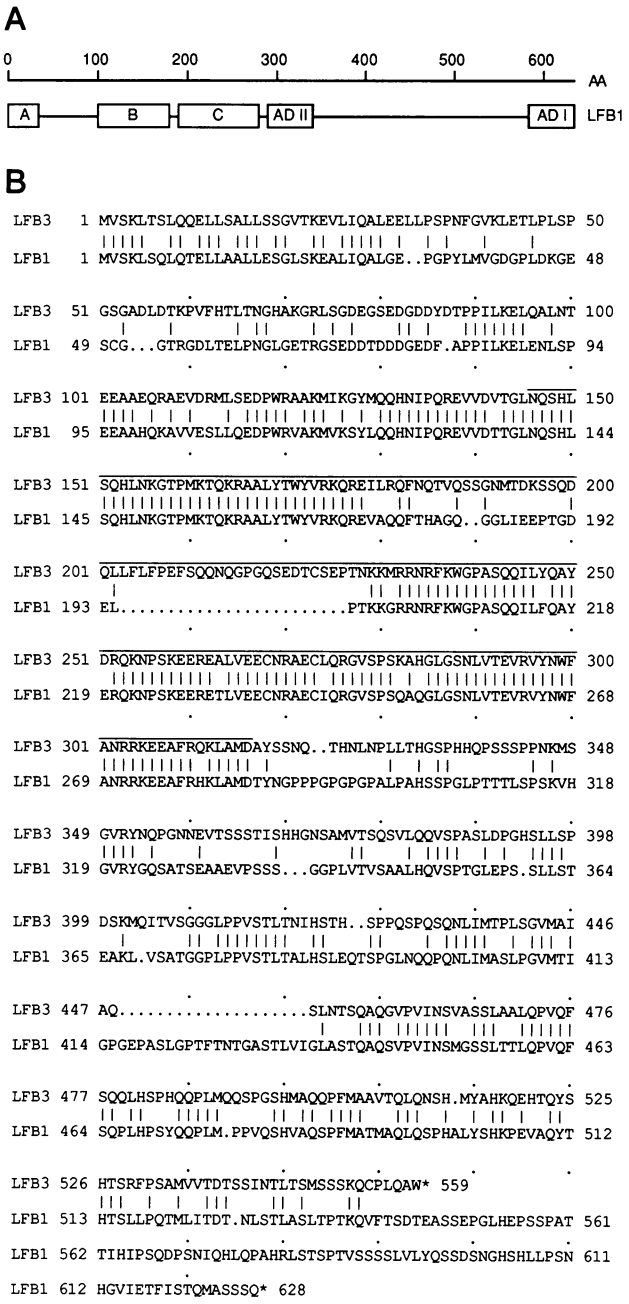

Fig. 2. Comparison of the predicted amino acid sequence of LFB3 with that of LFB1. (**A**). The functional organization of the LFB1 protein, as defined by Nicosia *et al.*, 1990. A = the dimerization domain (amino acids 1−33); B = the pseudo POU-specific domain (amino acids 100−184); C = extra-large homeodomain (amino acids 198−281); ADII and ADI = the two activation domains (amino acids 281−318 and amino acids 546−628 respectively). (**B**) Alignment of the mouse LFB3 and rat LFB1 amino acid sequences. The continuous line indicates the part of the rat LFB3 cDNA which has been sequenced. It displays no amino acid differences and only 12 (silent) nucleotide substitutions compared to the corresponding mouse cDNA region.

obtained with a specific combination of these oligonucleotides revealed, in addition to the predicted 466 bp fragment for LFB1, a slower migrating band weakly hybridizing to the LFB1 probe. This extra band was observed in a variety of cell lines like C2 and H5 cells, where the LFB1 specific band was absent, as well as in early embryos and in differentiated F9 murine embryonic carcinoma cells (data not shown). We

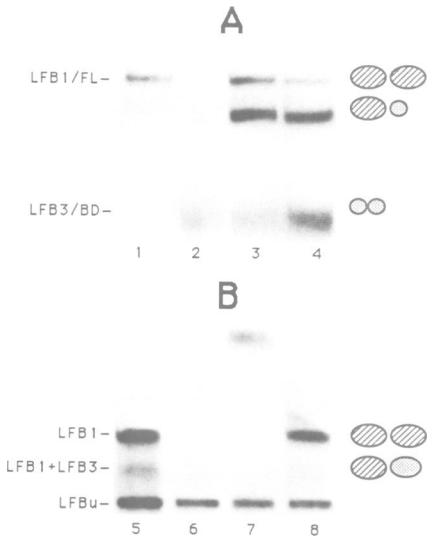

Fig. 3. Formation of LFB3/LFB1 heterodimers *in vitro* and *in vivo*. (**A**) Band-shift experiment showing the formation of heterodimers between *in vitro* translated LFB1/FL and LFB3/BD. 1 = LFB1/FL alone; 2 = LFB3/BD alone. 3 and 4 = LFB1-FL + LFB3-BD, mixed at a 2/1 and 1/2 ratio respectively. (**B**) Band-shift with Hep3B nuclear extracts. 5 = band-shift in the absence of antibodies; 6 = + αB1-A; 7 = + αB1-B; 8 = + αB3. OD scanning of lanes 5 and 8 shows a 4-fold reduction of LFB1+LFB3 relative to LFB1 band upon addition of aB3.

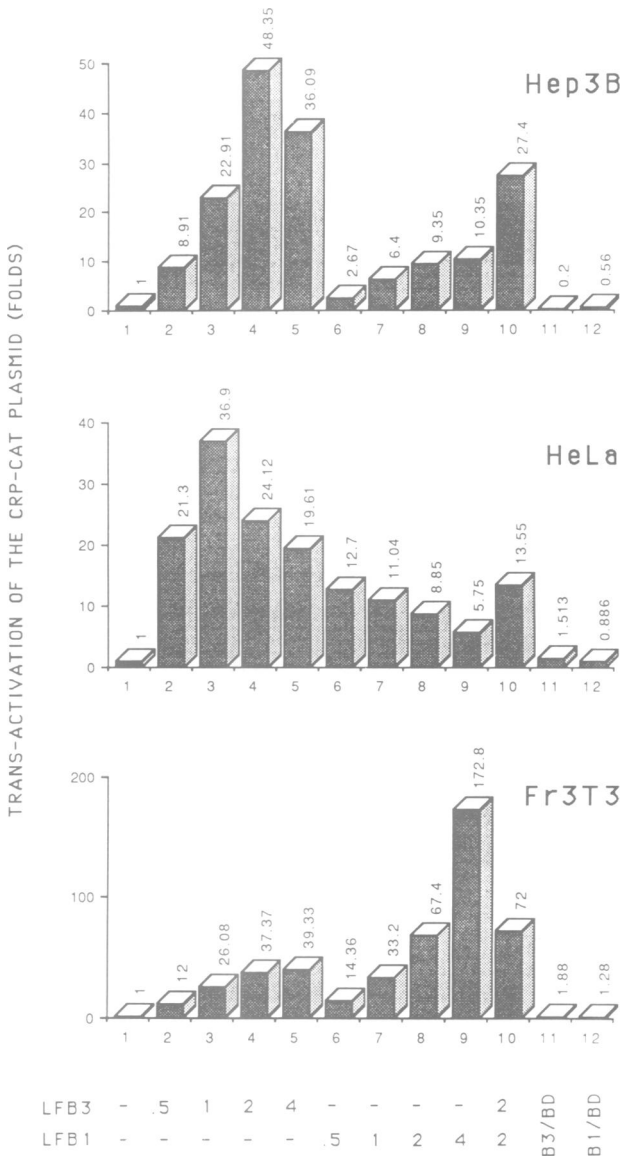

Fig. 4. Transactivation of an LFB1-dependent promoter by LFB3 and LFB1 in different cell lines. Values on the the *Y* axis = fold of transactivation, obtained dividing the CAT activity (normalized) of each sample by that of the target gene alone (lane 1) (see the Materials and methods section for a description of the normalization procedure). The amounts (μg) of the transactivating H2-LFB3 and H2-LFB1 plasmids are indicated at the bottom of the Figure. Samples in lanes 11 and 12 have been transfected with 4 μg of H2-LFB3/BD (11) or H2-LFB1/BD (12) respectively.

cloned and sequenced this segment of amplified cDNA from both rat and mouse. Using this cloned fragment as a probe, we isolated the full length cDNA from a mouse differentiated F9 cell cDNA library. The DNA sequence obtained is clearly different from that of LFB1, but the predicted amino acid sequence shows several significant homologies with LFB1 (Figure 2). In particular, there are three regions (amino acids 1−32, 89−178 and 229−318) which display a strong homology to LFB1. Interestingly, these conserved regions correspond to the A (dimerization domain), B (pseudo POU-specific domain), and C (homeodomain) elements which are essential for the DNA binding activity of LFB1. In contrast, neither of the two transcriptional activation domains identified in LFB1 by *in vitro* transcription assays (Nicosia *et al.*, 1990) are found in this new LFB1-like gene.

Polyclonal antibodies (αB3) were raised against the highly divergent region between amino acids 179 and 228. Addition of αB3 to band-shift reactions with H5 extracts eliminates LFB3−DNA interaction (Figure 1, lane 8), without interfering with LFB1 nor with LFBu (Figure 1, lane 4). On this basis, we conclude that this LFB1-like gene encodes the LFB3 DNA binding protein. Interestingly, αB3 inhibits the formation of the H complex in band-shift interference experiments with H4II extract (Figure 1, lane 4), suggesting that the H complex must contain, in addition to LFB1, LFB3 or an immunologically related protein.

The recombinant LFB3 protein binds as a dimer and forms heterodimers with LFB1 in vitro and in vivo

The sequence similarity with LFB1 in the first 32 amino acids suggested that LFB3, like LFB1, might bind to the target DNA as a dimer, and that it might even form heterodimers with LFB1. We performed band-shift experiments with *in vitro* translated recombinant LFB3, showing that it binds to the LFB1 recognition site; a deletion mutant containing only the N-terminal part of LFB3 (LFB3/BD, from amino acids 1 to 315) still binds to DNA (Figure 3, lane 2). LFB3 binds to DNA as a dimer, as shown by mixing the full length (LFB3/FL) with LFB3/BD (data not shown). Similarly, by mixing a full length LFB1 (LFB1/FL) and LFB3/BD *in vitro* translated proteins we found, in addition to the complexes characteristic of the two homodimers, a band of intermediate mobility corresponding to the LFB1/FL + LFB3/BD heterodimer (Figure 3, lanes 3 and 4). The reciprocal combination (LFB3/FL versus LFB1/BD) and the combination of the two full length proteins gave identical results (data not shown). It must be noted that mouse LFB3 and rat LFB1 proteins have been used for these dimerization experiments.

Band shift experiments with nuclear extracts from the Hep3B human hepatoma cell line revealed three retarded bands (Figure 3, lane 5). On the basis of the migration in the gel and the response to the three different antibodies, the upper and the lower band could be identified respectively as LFB1 and LFBu. The intermediate band is competed by all the three antibodies, suggesting that it must contain both LFB3 and LFB1. We interpret this band-shift as the binding of an LFB3/LFB1 heterodimer on the target DNA. We have observed an analogous intermediate complex in H4II (H in Figure 1, lane 1), in C2rev7, in liver and in kidney nuclear extracts, in which both LFB3 and LFB1 are present (data not shown).

LFB3 stimulates transcription from a LFB1-dependent promoter in transient expression experiments

LFB1 is a transcriptional activator, whose activity *in vitro* depends on the integrity of two domains, ADI and ADII (Nicosia *et al.*, 1990). The LFB3 C-terminal region shows no homology with either ADI or ADII. To test the transcriptional activity of LFB3 we performed cotransfection experiments in different cell lines. LFB3 was expressed under the control of the Major Histocompatibility Complex H2 gene promoter (H2-LFB3) (Kimura *et al.*, 1986). As a reporter we used a plasmid containing the human C-reactive protein promoter fused to the bacterial chloramphenicol-transacetylase reporter gene (CRP-CAT), since this construct is highly responsive to LFB1 in cotransfection experiments (Toniatti *et al.*, 1990). These constructs were cotransfected into Hep3B, which displays endogenous levels of LFB3 and LFB1, or into HeLa and Fr3T3 cell lines which contain neither of the two proteins.

The results of these experiments, reported in Figure 4, show that the activity of the reporter gene was greatly increased by cotransfection with increasing amounts of H2-LFB3 plasmid (compare lane 1 with lanes 2−5) in all the three cell lines. This effect was similar to that observed with H2-LFB1 (lanes 6−9). As a control, a plasmid expressing only the LFB3 binding domain (H2-LFB3/BD, amino acids 1−315) was not able to transactivate the CRP-CAT construct (lane 11). We conclude that the C-terminal portion of the LFB3 protein must contain a transcriptional activation domain(s), although different from that of LFB1.

The finding that LFB3 forms heterodimers with LFB1 raises the question of a possible role of these heterodimers in the regulation of transcription. Cotransfection of LFB3 and LFB1 into the same cells yielded levels of activation which are approximately an average of the two independent effects (compare lane 10 with the average of lanes 5 and 9). We conclude that, within the limitations of this experimental approach, there is no clear cooperative effect which could be ascribed to a putative heterodimer.

Interestingly, the relative efficiency of the two trans-activators varies among the different cell lines: LFB3 is more active than LFB1 in Hep3B and HeLa cells, while the reverse is true in Fr3T3 cells. Furthermore, the dose response profiles to the two transactivators in the three cell lines are different: in HeLa and, at higher doses, in Hep3B cells, we observed a progressive reduction of the transactivation efficiency with increasing amounts of the cotransfectant plasmid, probably due to a 'squelching' phenomenon. A possible explanation of these findings is that LFB3 and LFB1 activation domains interact with different transcriptional

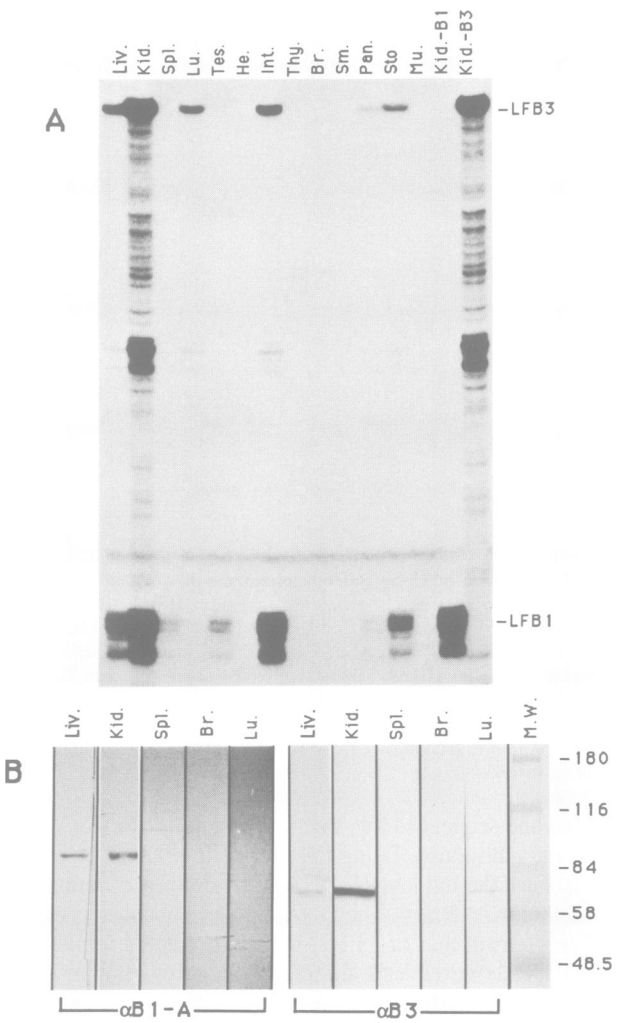

Fig. 5. LFB3 and LFB1 distribution in different adult rat tissues. (**A**) RNase mapping. Liv. = liver; Kid. = kidney, Spl. = spleen;, Lu. = lungs; Tes = testis; He. = heart; Int. = small intestine; Thy. = thymus; Br. = brain; Sm. = sub-maxillary glands; Pan. = pancreas; Sto. = stomach; Mu. = muscle. Kid.-B1 and Kid.-B3 are kidney samples showing the pattern obtained only with LFB3 or LFB1 probe respectively. (**B**) Western blot analysis. The mol. wts of the marker bands are expressed in kd.

'adaptors', which are present in different amounts in the three cell lines (Ptashne and Gann, 1990).

LFB3 and LFB1 are expressed in several specialized epithelia

LFB1 has been identified as a *trans*-acting factor regulating the expression of liver-specific genes. However, it has been reported that LFB1 mRNA is present in tissues other than liver, like kidney and intestine (Baumhueter *et al.*, 1990). We have studied the expression of LFB3 and LFB1 in various tissues by RNase mapping. In these experiments, an LFB3 probe covering the homeodomain yields three bands, the higher and more abundant one corresponding to the expected protected region (Figure 5A). The two shorter bands could be generated by unspliced precursors or by alternatively spliced products, since their lengths approximately correspond to the distance between the 3′ end of the probe and an intron−exon junction of the LFB3 gene (our unpublished results). The LFB1 probe yields multiple

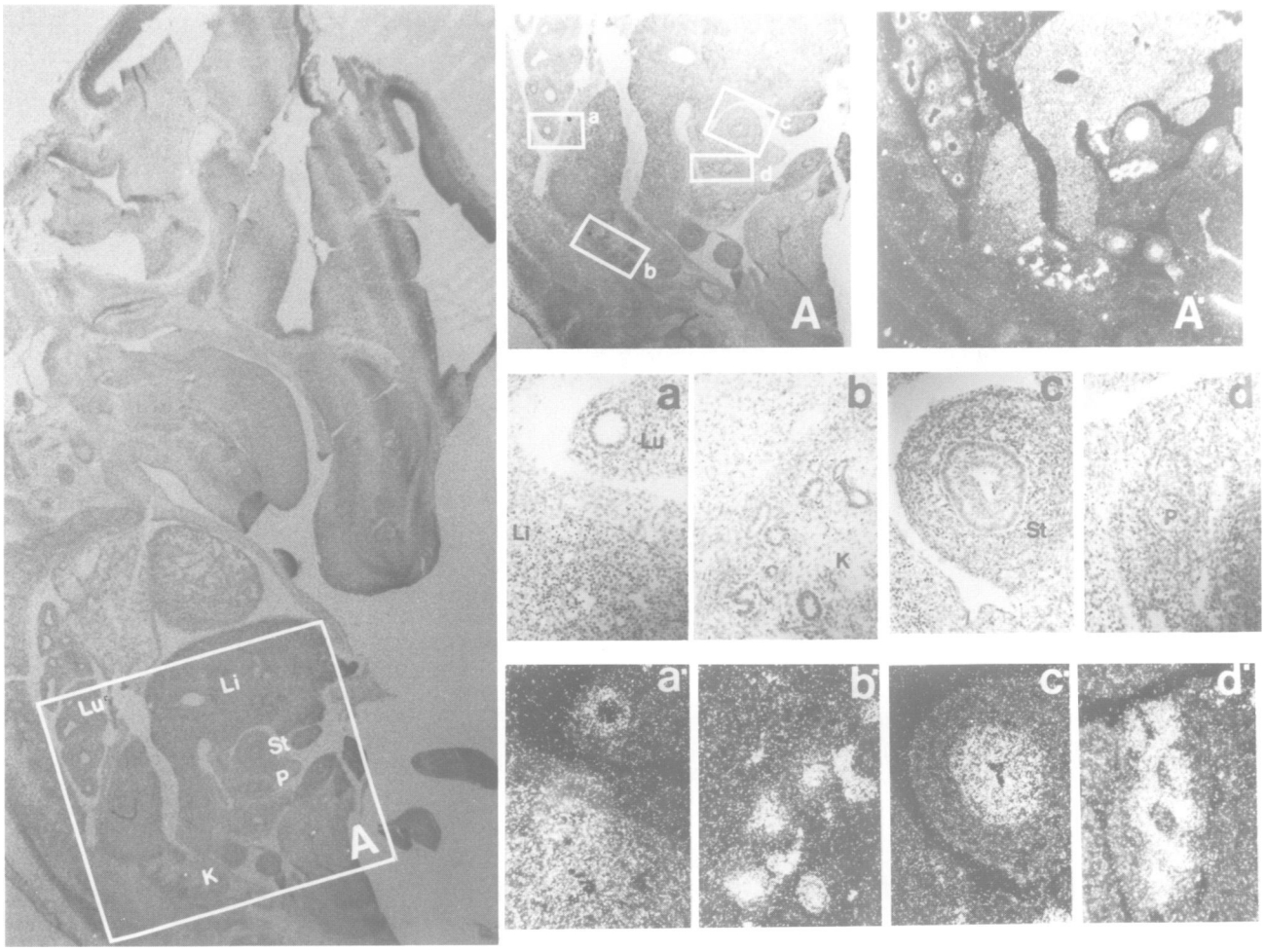

Fig. 6. *In situ* hybridization with LFB3 anti-sense probe on 14.5 days rat embryo sections.
Lu = lungs; Li = liver; St = stomach; P = pancreas; K = kidney. A, a, b, c, d = toluidine blue stained sections; A′, a′, b′, c′, d′ = autoradiography of the corresponding sections.

bands clustered around the expected position for the LFB1 transcript. The results of these experiments (Figure 5A) indicate that both LFB1 and LFB3 mRNAs are found mainly in liver, kidney, pancreas and in the digestive tract. However, there are some organs, like lungs, that contain only LFB3, and others, like spleen and testis, that contain only small amounts of LFB1 mRNA. We have also detected LFB3 mRNA, but not LFB1, in esophagus and thyroid (data not shown). RNase mapping data are not really quantitative and do not allow an accurate measurement of the absolute amount of mRNA; nevertheless, they allow a thorough determination of the relative abundance of LFB1 and LFB3 transcripts. However, these data are partially confirmed by Western analysis: we have detected LFB3 in liver, kidney and lungs, while LFB1 was found only in the first two tissues (Figure 5B).

We have performed *in situ* hybridization experiments with 14.5 days rat embryos. The results obtained with the LFB3 probe are shown in Figure 6. LFB3 transcripts are detected in liver, kidney, lungs, intestine, stomach and pancreas. More interestingly, they are localized only in the epithelial component of these tissues. *In situ* hybridization with a LFB1 probe yielded identical results for liver, intestine, stomach and kidney. There was no LFB1 signal in lungs, in agreement with the RNase mapping results; however, we could not obtain a clear signal above background levels in pancreas with the LFB1 probe. In Figure 7 we show an enlargement of the liver from a 15.5 days rat embryo, showing details of a perisinusoidal region. It is clear that there is a difference in the tissutal distribution of LFB3 and LFB1: LFB3 transcripts are more concentrated in the perisinusoidal hepatocytes than in the rest of the liver parenchyma, while LFB1 transcripts are almost uniformly distributed. Differences in the distribution of the two transcripts have also been observed in kidney (Lazzaro *et al.*, unpublished results).

LFB3 and LFB1 are differentially regulated in early embryonal stages, in F9 embryonal carcinoma cells and in rat hepatoma cell lines

As shown in the previous section, the expression pattern of LFB3 in adult tissues, as well as its tissutal localization, differs from that of LFB1. Preliminary PCR experiments suggested that the two genes are also differentially regulated during development. RNase mapping experiments using mouse embryo RNAs show that LFB3 transcripts are already present at day 6.5 (early egg cylinder stage) whereas barely detectable amounts of LFB1 mRNA appear only at day 9.5, corresponding to the formation of the yolk sac (Figure 8, first panel). Similarly, in rat embryos, LFB3 mRNA is

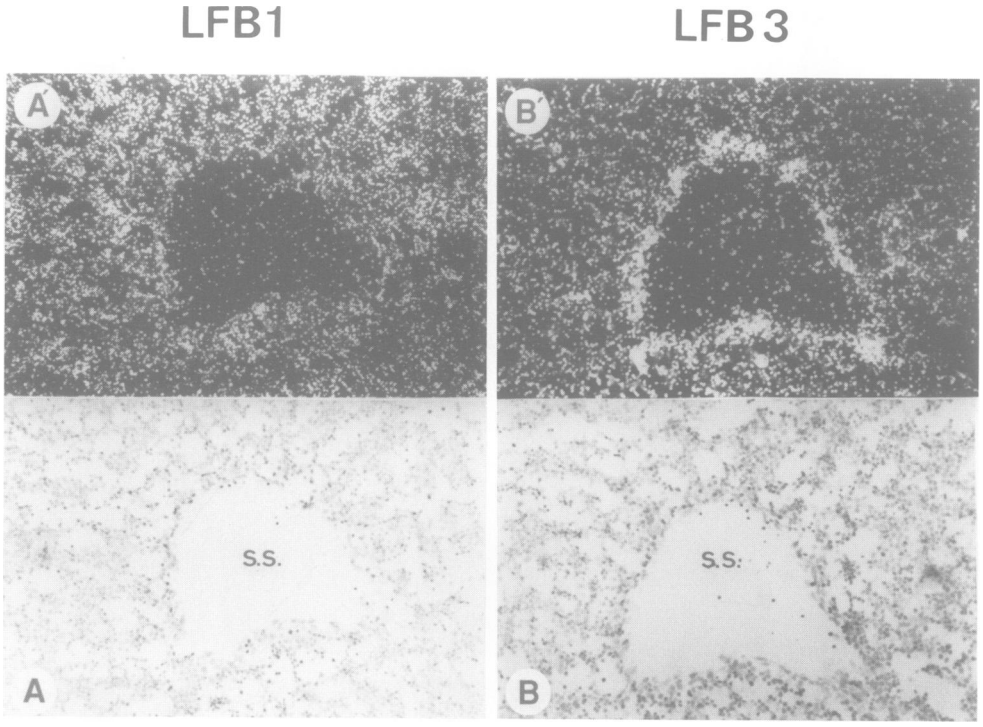

Fig. 7. *In situ* hybridization with LFB3 and LFB1 anti-sense probe on rat embryonic liver sections. A, B = toluidine blue stained sections; A', B' = autoradiography of the corresponding sections. S.S. = sinusoidal space.

present already at day 9, whereas LFB1 appears later and gives a strong signal only at day 12, mainly in the yolk sac and in the embryonic liver (Figure 8, second panel). LFB3 and LFB1 are differentially expressed in rat hepatoma cell lines. RNase mapping experiments reveal LFB3 transcripts both in dedifferentiated (C2 and H5) and differentiated (H4 and C2rev7) hepatoma cells. In contrast, LFB1 is present only in differentiated cells (Figure 8, third panel). Analogous results have been obtained by Western blot analysis (data not shown). Finally, in F9 embryonal carcinoma cells, often used as a model system for early developmental events, LFB3, but not LFB1 is induced by retinoic acid (Figure 8, fourth panel). All these data indicate that LFB3 transcription precedes that of LFB1.

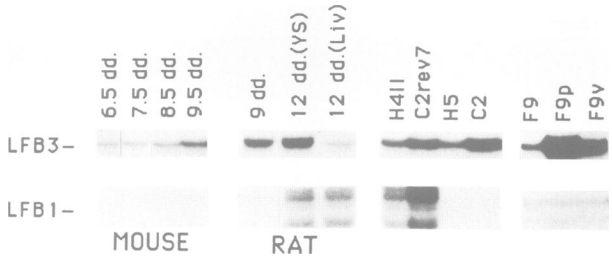

Fig. 8. RNase mapping of LFB3 and LFB1 transcripts in different embryonal stages and cell lines. 6.5, 7.5, 8.5 and 9.5 dd. = mouse embryo samples, days post-coitum (PC). 9 dd = 9 days (PC) rat embryo. 12 dd.(YS) and 12 dd.(Liv) = respectively the yolk sac and the embryonic liver dissected from a 12 days old rat embryo. F9 = undifferentiated murine F9 teratocarcinoma cell line; F9p = differentiated 'parietal-like' F9 cultures; F9v = differentiated 'visceral-like' F9 cultures. H4II, H5, C2, C2rev7 = differentiated (H4II, C2rev7) and dedifferentiated (H5, C2) rat hepatoma cell lines.

Discussion

LFB3, a new member of an extra-large homeodomain protein family, forms heterodimers with LFB1

The functional mapping of the LFB1 domains which are essential for DNA binding and for transcriptional activation has been reported (Nicosia *et al.*, 1990). Efficient and specific DNA binding requires the integrity of three LFB1 regions: the A domain, mediating dimerization; the B domain, important for the specificity of DNA recognition (R.De Francesco and L.Tomei, personal communication) and the C domain, which contains an extra-large homeodomain. The same protein elements (with some amino acid changes) are found in the LFB3 DNA binding domain. The most conserved region between LFB3 and LFB1 is the homeodomain, with 91% homology compared with the 71 and 79% of the A and B domains respectively. LFB3 and LFB1 thus belong to a new homeoprotein family, characterized by an extra-large homeodomain and by the capacity to dimerize and heterodimerize. No heterodimer formation between homeoproteins has been previously observed, excepted for yeast mating type factors (Dranginis, 1990).

We have shown that LFB3 can form heterodimers with LFB1 both *in vitro* and *in vivo*. What is the role played by the LFB3/LFB1 association? An increase in affinity for the DNA target sequence has been observed in the case of c-jun and c-fos heterodimers (reviewed by Busch and Sassone-Corsi, 1990). In another case, heterodimers between the *Drosophila* genes *daughterless* and the T3 protein of the *achaete–scute* complex give rise to changes in the sequence specificity (Jones, 1990). However, in our experimental system, we have observed no difference in the DNA binding activity of the LFB3/LFB1 heterodimer compared with that of the two homodimers.

Another possibility is that the heterodimeric complex might acquire new transcriptional activation properties. LFB3 and LFB1 diverge in the C-terminal region, especially in the LFB1 transcription activation domains ADI and ADII. Nevertheless, LFB3 is at least as potent a transcriptional activator as LFB1. Its activation function must be contained in the C-terminal half of the protein (from amino acids 316–528) since a deletion mutant, which contains only the first 315 amino acids, fails to transactivate the target promoter in our transfection assay. LFB3 and LFB1 display a significant difference in their activation potential: LFB3 seems to be more efficient than LFB1 in Hep3B and in HeLa, and less efficient in Fr3T3. Since they are likely to interact with different components of the transcriptional machinery, for instance with different 'adaptors', differential activation may reflect the concentration and the tissue distribution of these adaptors. The difference in the activation domains of LFB3 and LFB1 may result in the LFB3/LFB1 heterodimer performing a specific function, for example making the complex sensitive to a given regulatory factor acting via protein–protein interactions. We have recently described (Poli *et al.* 1990) that heterodimer formation between IL6DBP, a relatively weak transactivator inducible by interleukin 6 (IL6), and C/EBP, a strong transactivator not responsive to IL6, results in the high activation potency of C/EBP combined with the IL6 inducibility of IL6DBP.

LFB3 and LFB1 are differentially regulated

During mouse development, LFB3 transcription is already detectable at day 6.5 whereas LFB1 mRNA is found only at day 9.5. Therefore, LFB3 is expressed at an earlier stage than any of the so far identified genes which carry the cognate DNA recognition site. In contrast, transcription of LFB1 is concomitant with the expression of the genes to whose promoters it binds. However, we cannot exclude the possibility that LFB3, alone or in association with other proteins, binds to additional sites perhaps in promoters of genes expressed early during development.

LFB3, but not LFB1, is present in dedifferentiated hepatoma cell lines (C2 and H5). Assuming that during dedifferentiation cells revert to a stage which is similar to an earlier step of the physiological maturation process, this observation further supports the idea that LFB3 is expressed earlier than LFB1. In F9 cells LFB3 and LFB1 are also differentially regulated. These cells can be induced to differentiate into endodermal-like cultures by retinoic acid (RA) treatment; upon RA stimulation, LFB3 (but not LFB1) is strongly induced. It has been reported that a set of human homeobox containing genes are sequentially activated by RA according to their temporal expression *in vivo* (Simeone *et al.*, 1990). It is tempting to speculate that also the inducibility of LFB3 by RA is a feature related to its earlier appearance during development.

The localization of LFB3 expression in the perisinusoidal hepatocytes could reflect a response to environmental cues, or perhaps to morphogenetic gradients. A similar perisinusoidal distribution has been described for the transcripts of some liver-specific genes, like glutamine synthetase, leading to the hypothesis that the periportal hepatocytes could represent a functionally (and biochemically) distinct population (Gebhardt *et al.*, 1988). The possibility that LFB3 could play a role in this hepatocyte population should be taken into account.

Relationship between LFB3 and vHNF1

The vHNF1 DNA binding activity was originally observed in dedifferentiated hepatoma cell lines and in fibroblast–hepatoma cell hybrids, which have lost the ability to express several liver-specific functions and do not transcribe LFB1 (Cereghini *et al*, 1988; Baumhueter *et al.*, 1988; Cereghini *et al.*, 1990). The involvement of vHNF1 in the extinction phenomenon, or at least in repression of liver-specific gene expression was postulated. We present results indicating that the original vHNF1 activity actually consists of two distinct components. One of these activities, LFB3, is a transcriptional activator and corresponds to the gene which we have cloned and characterized. The second activity, LFBu, is a ubiquitous protein whose role has not been further investigated.

On the basis of the transactivation results presented in this paper, the hypothesis that LFB3 is a repressor seems unlikely. Moreover, the original observation that the presence of vHNF1 and LFB1 are mutually exclusive clearly does not hold, since LFB3 is also expressed in differentiated cells and in the liver itself. The failure to reveal the LFB3 homodimer binding activity in differentiated hepatoma cell lines and in liver could be due to a recruitment of all the LFB3 protein into heterodimers with LFB1, which is present in higher amounts.

The role of LFB3 and LFB1 in liver-specific gene expression

Work in many laboratories has led to the conclusion that the integrity of the LFB1 binding site is necessary for efficient transcription of many liver-specific promoters *in vitro*, in cultured cells and in transgenic mice (for a review see De Simone and Cortese, 1988). However, since LFB1 and LFB3 are present in several other cell types, we must conclude that their activity is not sufficient to determine transcription of liver-specific genes. A similar situation has been found for many other putative tissue-specific *trans*-acting factors. The octamer binding protein oct-2 is also apparently expressed in cells other than lymphocytes (Cockerill and Klinken, 1990); the pituitary *trans*-acting factor Pit-1 is also expressed in brain (Sharp and Cao, 1990); TTF1, the recently characterized thyroid specific factor, is also expressed in lungs (Guazzi *et al.*, 1990); Eryf-1, the erythropoietic specific factor, is found also in myelocytes (Romeo *et al.*, 1990). The current idea is that transcription factors have a wider domain of expression than a single cell type: whenever different domains overlap, a unique combination of *trans*-acting factors is generated, which is characteristic of a cell type and which causes cell type specific transcription. In this ideal transcriptional Venn diagram, the hepatocyte phenotype is the result of the intersection of the domain of expression of several transcription factors, including LFB1/HNF1, LFB3, LFA1 (Hardon *et al.*, 1988), C/EBP (Landschulz *et al.*, 1988), IL6DBP/LAP (Poli *et al.*, 1990), DBP (Mueller *et al.*, 1990) and HNF3 (Lai *et al.*, 1990). For all of these factors a role in liver-specific transcription has been established or suggested. What might be their role in other cell types? In the case of C/EBP McKnight and co-workers have suggested that it might regulate transcription of genes involved in energy metabolism (McKnight *et al.*, 1989). These genes are expressed in various cell types which, taken together, represent the C/EBP expression domain.

LFB3 and LFB1 are expressed in several tissues which are not related by an embryonic lineage or by topology. The only common feature of these tissues is to have a specialized epithelium, and *in situ* hybridization experiments show clearly that LFB3 and LFB1 are expressed in the polarized epithelia. Epithelial cells could be then the expression domain of LFB3 and LFB1. In these cells, specific proteins are present, for example cytokeratins, uvomorulin, cingulin, laminin A (Rodriguez-Boulan and Nelson, 1989). The coordinated expression of genes coding for these proteins might be the result of the implementation of an 'epithelial' genetic program, in which both LFB3 and LFB1 might be involved, the first at an earlier stage than the second. Support for the notion of an 'epithelial program' derives from the recent observation that the mutation *crumbs* in *D.melanogaster* specifically blocks the development of epithelial tissues (Tepass *et al.*, 1990). The role of LFB3 and/or LFB1 in the epithelial differentiation in mammalian cells will be the subject for further investigation.

Materials and methods

Nuclear extracts, gel retardation assays and antibody competition

Protein extracts from isolated nuclei of cultured cells were prepared according to Lee *et al.* (1988). Tissue nuclear extracts were prepared according to Lichtsteiner *et al.* (1987). Gel retardation experiments with crude nuclear extracts were performed incubating 2 μg of nuclear proteins in a buffer containing 60 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 0.75 mM DTT, 7.5% (v/v) glycerol, 0.05% Nonidet-P40, 3 μg BSA and 1 μg of poly(dI−dC) as carrier DNA. After 15 min on ice, ~1 ng (20 000 c.p.m.) of end-labeled double-stranded oligonucleotide was added to the reaction mixture and incubated for a further 30 min on ice. 5 μl of 20% Ficoll were added and the samples were fractionated on a 6% acrylamide gel (30:1 acrylamide/bisacrylamide, 0.25×TBE) at 10 V/cm at 4°C. Band-shift assays with *in vitro* translated proteins were performed in a similar way by using 5 μl of *in vitro* translation reaction. As a probe for these experiments a double-stranded oligonucleotide containing the human α1-antitrypsin LFB1 binding site has been used (B1 oligonucleotide; Monaci *et al.*, 1988). For the analysis of the dimerization, recombinant forms of proteins were either co-translated or mixed after translation. The mixed proteins were incubated for 15 min at room temperature or at 37°C followed by a further incubation for 30 min on ice, and used for the gel retardation assay. Antibody-interference with band shift were performed by incubating the protein extract with 1 μl antiserum (1:10 dilution) for 30 min on ice.

Antisera and western blotting

The αB1-A and αB3 polyclonal antibodies were raised against respectively the amino acids 1−287 region of LFB1 or the amino acids 179−228 of LFB3, fused to the MS2 phage polymerase in a derivative of the pEx29 expression vector (Klinkert *et al.*, 1985). αB1-B antibodies were raised against a synthetic peptide corresponding to the C-terminal region of LFB1 (amino acids 595−628). The three polyclonal antibodies were raised into rabbits. Boost injections were done at 2 week intervals, and the sera used for the experiments described here were obtained after three boosts. The antibodies were tested for cross-reactivity against the recombinant proteins.

For Western blot analysis, nuclear proteins were first subjected to electrophoresis (50−150 μg protein/lane) in 7.5% polyacrylamide gels (Laemmli, 1970), containing 15% sucrose. We used a Genie electrophoretic blotter (IDEA scientific company, USA) according to the instructions of the supplier with a Tris−glycine blotting buffer. The proteins were blotted for 30 min to 1 h onto nitrocellulose filters. After blotting the filters were stained with Ponceau Red solution (Serva). Before starting the antibody reaction the filters were incubated in 'Blotto' (10% dry non-fat milk, 1×PBS, 0.1% Triton, 0.1% NaN_3) at 4°C for a minimum of 1 h. The antibodies were diluted 1:1000 in Blotto and the filters incubated in this solution for 1 h, washed 3×5 min with 1×PBS + 0.1% Triton and kept in Blotto for 15 min. The filters were then treated with an appropriate dilution in Blotto of a goat anti-rabbit antibody conjugated with alkaline phosphatase (Sigma) for 1 h. After washing again three times with 1×PBS + 0.1% Triton, the bands were visualized by incubating the filters in a solution containing 0.66% 4-nitrotetrazoliumchloride-blue-hydrate (in 100% dimethylformamide), 0.33% 5-bromo-4-chloro-3-indolylphosphate-*p*-toluidine salt (in 70% dimethylformamide), 100 mM Tris, pH 9.6, 100 mM NaCl and 5 mM $MgCl_2$ for 5−30 min. All steps of the antibody reaction were performed at room temperature.

cDNA synthesis, PCR amplification and cDNA library screening.

For each PCR assay, 2 μg of total RNA was used to synthesize the anti-sense cDNA strand with an LFB1 specific primer in a reverse transcription reaction (AMV-RT, Pharmacia). One tenth of this cDNA was used in a *Taq* polymerase PCR reaction (Perkin−Elmer Cetus), using the same oligonucleotide as before plus another LFB1-specific oligonucleotide able to 'prime' on the cDNA strand. Both reactions were carried out according to the indications of the supplier. The cDNA library was constructed in λgt10 vector (Huyn *et al.*, 1985) with mRNA from F9 cells cultured for 5 days in a medium containing retinoic acid (10^{-4} M).

In vitro transcription and translation

In vitro transcriptions were performed using 1 μg of linearized template in a 10 μl reaction containing 40 mM Tris−HCl (pH 8.0), 50 mM NaCl, 8 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 0.5 mM rNTPs, 10 U RNasin, 0.27 U m^7GpppG and 12.5 U of T7 RNA polymerase (Stratagene). The mixture was incubated for 1 h at 37°C, diluted to 100 μl with water and extracted with phenol/chloroform/isoamyl-alcohol. The translated RNA was purified through a 1 ml Sephadex G50 spun column, ethanol precipitated and resuspended in 10 μl of water. Antisense riboprobes for RNase mapping and *in situ* hybridization were prepared using the T3 RNA polymerase with the same procedure, but without m7GpppG and adding 15 μCi of [α-^{32}P]UTP (RNase mapping) or [α-^{35}S] UTP (*in situ* hybridization) (Toth *et al.*, 1987).

The *in vitro* translation reactions were carried out for 2 h at 25°C in a 50 μl reaction containing 1 μl of the *in vitro* transcribed RNA, 80 mM KOAc, 0.3 mM $MgCl_2$, 80 μM amino acid mix (minus methyonine), 37.5 μCi [^{35}S]methionine (Amersham SJ515) and 25 μl wheatgerm extract (Promega). The ^{35}S-labeled translation products were analyzed on SDS−PAGE.

Plasmid constructions

The full length cDNA of LFB3 and a fragment containing the amino acids 1−315 region (LFB3/BD) were cloned in the bluescript SK+ (BSC) vector (Promega) in the T7 → T3 orientation for *in vitro* translation and expression of the protein. The BSC-LFB3/BD plasmid was also used to generate riboprobes for RNase mapping and *in situ* hybridization with T3 RNA polymerase. The same fragments were also cloned in the pH2 expression vector (Ruether, personal communication), generating H2-LFB3/FL and H2-LFB3/BD plasmids. Corresponding LFB1 plasmids expressing either the full length (H2-LFB1/FL) or the amino acids 1−287 (H2-LFB1/BD) protein were also constructed.

Cell cultures, DNA transfections and CAT assays

Hep3B, HeLa, Fr3T3 and F9 cell lines were cultured in Dulbecco's modified Eagle's medium + 10% fetal calf serum at 5% CO_2; H4II, H5, C2 and C2rev7 were cultured in a 1:1 mixture of F12/NCTC-135 media (Gibco/Flow) + 10% fetal calf serum at 7% CO_2. F9 cells were induced to differentiate in 'parietal-like' cultures by growth in the presence of 10^{-4} M retinoic acid + 10^{-3} M dibutyryl-cAMP for 5 days. F9 'visceral-like' cultures were obtained by growing the cells in suspension and in the presence of 10^{-5} M retinoic acid.

DNA transfections were performed with the calcium phosphate precipitation technique as previously reported (De Simone *et al.*, 1987). Corresponding samples of the three cell lines were transfected with an aliquot of the same precipitate, and each precipitate was prepared in duplicate. For each 5 cm diameter plate the precipitate contained: (i) 4 μg of the target plasmid (−219CRP-CAT); (ii) variable amounts of the plasmids expressing LFB3 (H2-LFB3) or LFB1 (H2-LFB1). In each case, the total concentration of the expression plasmid was adjusted to 6 μg with the parental pH2 plasmid; (iii) 0.2 μg of Rous sarcoma virus−luciferase plasmid (de Wet *et al.*, 1987) as an internal control for transfection efficiency. The Rous sarcoma virus−CAT plasmid (Gorman *et al.*, 1982) was used as a control of the efficiency of transfection. Cells were harvested 48 h after transfection; CAT activity was assayed according to Gorman *et al.* (1982) and normalized to the luciferase activity of the same sample measured as described by de Wet *et al.* (1987). The transactivation was expressed as the ratio of the normalized CAT activity of the sample to that of the target construct alone. The mean values for each couple of samples have been plotted.

RNA extraction and RNase mapping analysis

Total RNA was extracted from cell cultures and tissues by using the guanidine-thiocyanate/acid phenol method (Chomczynski and Sacchi, 1987). Fresh tissues were homogenized directly in GTC buffer by an engine-operated homogenizer; cultured cells and early embryonic samples were resuspended in GTC buffer and disrupted manually by 3–4-fold aspiration through an extra-thin needle. Mouse embryonic samples were prepared by manually dissecting the uterus and collecting the implanted embryo along with a portion of the decidua. Rat embryonic samples at day 12 were prepared by micro-dissecting the embryo and collecting the yolk sac and the liver primordium.

In RNase mapping experiments, 30–50 μg of total RNA were annealed with 1.5×10^5 c.p.m. of the specific riboprobe at 45°C O/N, then digested with a mixture of RNase A + T1 and separated on a 6% sequencing gel. The rat and mouse LFB3 riboprobes were both 614 bp long, yielding a 510 bp protected band (amino acids 146–316); the rat LFB1 riboprobe was 308 bp, yielding a 224 bp protected band (amino acids 207–282), while the mouse LFB1 riboprobe was 191 bp, yielding a 156 bp protected band (amino acids 140–192).

In situ hybridization

In situ hybridization was performed according to a modification of the procedure described by Pintar and Lugo (1987). Rat embryos were fixed in phosphate buffered saline (PBS) + 4% paraformaldehyde at 4°C for 8 h, transferred in PBS + 15% sucrose at 4°C for 24 h, embedded in OCT compound (Miles laboratories), and stored frozen at −80°C. Hybridization and washes were performed according to Toth *et al.* (1987). The slides were coated with a 1:1 dilution of Kodak NTB2 emulsion, exposed for 3–4 weeks at 4°C, developed with Kodak D19 and Rapid Fix and stained with toluidine blue.

Acknowledgements

We wish to thank Drs I.Mattaj, M.Price, R.De Francesco, D.Ramji, L.Tomei and F.Tronche for critically reading the manuscript; G.Cilberto for the gift of the CRP-CAT plasmid, M De Felice for the gift of several rat RNAs and U.Ruether for the pH2 vector. This work has been partially supported by the EEC grant BAP-0115-D.

References

Arias,I.M., Popper,H., Schachter,D. and Shafritz,M.D. (1982) *The Liver, Biology and Pathobiology*. Raven Press, New York.

Baumhueter S.,Courtois,G. and Crabtree,G. (1988) *EMBO J.*, **7**, 2485–2493.

Baumhueter,S., Mendel,D. B., Conley,P. B., Kuo,C., J., Turk,C., Graves,M.K., Edwards,C. A. Courtois,G. and Crabtree,G. (1990) *EMBO J.*, **4**, 372–379.

Birkenmeier,E.H., Gwynn,B., Howard,S., Jerry,J., Gordon,J.I., Landschulz,W.H. and McKnight,S.L. (1989) *Genes Dev.*, **3**, 1146–1156.

Busch,S.J. and Sassone-Corsi,P. (1990) *Trends Genet*, **6**, 36–40.

Cereghini,S., Blumenfeld,M. and Yaniv,M. (1988) *Genes Dev.*, **2**, 957–974.

Cereghini,S., Yaniv,M., and Cortese,R. (1990) *EMBO J.* **9**, 2257–2263.

Chomczynski,P., and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.

Cockerill,P.N., and Klinken,S.P. (1990) *Mol. Cell. Biol.*, **10**, 1293–1296.

Courtois,G., Morgan,J.G., Campbell,L.A., Fourel,G. and Crabtree,G.R. (1987) *Science*, **238**, 688–692.

Courtois,G., Baumhueter,S. and Crabtree,G.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7937–7941.

Deschatrette,J and Weiss,M.C. (1980) *Cell*, **19**, 1043–1051.

Descombes,P., Chojikier,M., Lichtsteiner,S., Falvey,E., and Schibler,U. (1990) *Genes Dev.*, **4**, 1541–1551.

De Simone,V. and Cortese,R. (1988) In McLean,N. (ed.), *Oxford Surveys on Eukaryotic Genes*, Oxford University Press, Oxford, Vol. 5, pp. 51–90.

De Simone,V., Ciliberto,G., Hardon,E., Paonessa,G., Palla,F., Lundberg,L. and Cortese,R. (1987) *EMBO J.*, **6**, 2759–2766.

de Wet,J.R., Wood,K.V., DeLuca,M., Helinski,D.R. and Subramani,S. (1987) *Mol. Cell. Biol.*, 7, 725–737.

Dranginis,A. M. (1990) *Nature*, **347**, 682–685.

Feuerman,M.H., Godbout,R., Ingram,R.S. and Tilghman,S.M. (1989) *Mol. Cell. Biol.*, **9**, 4204–4212.

Finney,M. (1990) *Cell*, **60**, 5–6.

Frain,M., Swart,G., Monaci,P., Nicosia,A., Stämpfli,S., Frank,R. and Cortese,R. (1989) *Cell*, **59**, 145–157.

Gebhardt,R., Ebert,A. and Bauer,G. (1988) *FEBS Lett.*, **241**, 89–93.

Gorman,C.M., Merlino,G.T., Willingham,M.C., Pastan,I. and Howard,B.H. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6777–6781.

Guazzi,S., Price,M., De Felice,M., Damante,G., Mattei,M.G., and Di Lauro,R. (1990) *EMBO J.*, **9**, 3631–3639.

Hardon,E.M., Frain,M., Paonessa,G. and Cortese,R. (1988) *EMBO J.*, **7**, 1711–1719.

Huyn,T.V., Young,R.A. and Davis ,R.W. (1985) In Glover,D.M. (ed.), *DNA Cloning–A Practical Approach*, IRL Press, Oxford, Vol. I, pp. 49–78.

Kimura,A.,Israel,A., Le Bail,O. and Kourilsky,P. (1986) *Cell*, **44**, 261–272.

Klinkert,M., Herrmann,R., and Schaller,H. (1985) *Infection Immun.*, **49**, 329–335.

Jones,N. (1990) *Cell*, **61**, 9–11.

Laemmli,U.K. (1970) *Nature*, **277**, 680–685.

Lai,E., Prezioso,V.R., Smith,E., Litvin,O., Costa,R.H., and Darnell,J.E. (1990) *Genes Dev.*, **4**, 1427–1436.

Landschulz,W.H., Johnson,P.F., Adashi,E.Y., Graves,B.J. and McKnight,S.L. (1988) *Genes Dev.*, **2**, 786–800.

Lichtsteiner,S. and Schibler,U. (1989) *Cell*, **57**, 1179–1187.

Lichtsteiner,S., Wuarin,J. and Schibler,U. (1987) *Cell*, **51**, 963–973.

McKnight,S.L., Lane,M.D. and Gluecksohn-Waelsch,S. (1989) *Genes Dev.*, **3**, 2021–2024.

Monaci,P., Nicosia,A. and Cortese,R. (1988) *EMBO J.*, **7**, 2075–2087.

Mueller,CR., Maire,P., and Schibler,U. (1990) *Cell*, **61**, 279–91

Nicosia,A., Monaci,P., Tomei,L., De Francesco,R., Nuzzo,M., Stunnenberg,H., and Cortese,R. (1990) *Cell*, **61**, 1225–1236.

Pintar,J.E. and Lugo,D.I. (1987) In Valentino,K.L. Eberwine,J.H. and Barches,J.D. (eds), *In Situ Hybridization Applications to Neurobiology*, Oxford University Press, Oxford pp. 179–196.

Poli,V., Mancini,F., and Cortese,R. (1990) *Cell*, **61**, in press.

Ptashne,M., and Gann,A.A.F. (1990) *Nature*, 346, 329–331.

Rodriguez-Boulan,E., and Nelson,W. J. (1989) *Science*, **245**, 718–725.

Romeo,P. H., Prandini,M. H., Joulin,V., Mignotte,V., Prenant,M., Vainchenker,W., Marguerie,G., and Uzan,G.(1990) *Nature*, **344**, 447–449.

Schorpp,M., Kugler,W., Wagner,U. and Ryffel,G.U. (1988) *J. Mol. Biol.*, **202**, 307–320.

Scott,M.P., Tamkun,J.W. and Hartzell,G.W.III (1989) *Biochim. Biophys. Acta*, **989**, 25–48.

Sharp,Z. D., and Cao,Z. (1990) *BioEssays*, **12**, 80–85.

Simeone,A., Acampora,D., Arcioni,L., Andrews,P.W., Boncinelli,E., and Mavilio,F. (1990) *Nature*, **346**, 763–766.

Tepass,U., Theres,C., and Knust,E. (1990) *Cell*, **61**, 787–799.

Toniatti,C., Demartis,A., Monaci,P., Nicosia,A., and Ciliberto,G. (1990) *EMBO J.*, 9, 4467–4475

Toth ,L.E., Shawin,K.L., Pintar,J.E. and Chi Nguyen-Huu,M. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6790–6794.

Tronche,F., Rollier,A., Bach,I., Weiss,M., and Yaniv,M. (1989) *Mol. Cell. Biol.*, **9**, 4759–4766.

Received on December 13, 1990; revised on January 22 1991.

Note added in proof

The nucleotide and protein sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ nucleotide sequence databases under the accession number X55842.