The EMBO Journal vol.7 no.6 pp.1711–1719, 1988

Two distinct factors interact with the promoter regions of several liver-specific genes

Elizabeth M.Hardon, Monique Frain[1], Giacomo Paonessa and Riccardo Cortese

European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, FRG
[1]On leave of absence from the Laboratoire d'Enzymologie du CNRS, 91190, Gif-sur-Yvette, France

Communicated by R.Cortese

A segment of the human α1-antitrypsin (α1AT) 5′-flanking region comprising nucleotides −137 to −37 from the start of transcription is sufficient to drive liver-specific transcription from the homologous α1AT promoter and from the heterologous SV40 promoter. In this paper we characterize two proteins, LF-A1 and LF-B1, whose ability to bind wild-type and mutant α1AT promoter segments correlates with the ability of these segments to activate transcription *in vivo*. DNase I protection and methylation interference analysis reveals that LF-A1 recognizes sequences present in the regulatory region of the human α1-antitrypsin, apolipoprotein A1 and haptoglobin-related genes. These sequences share a common 5′ TG$^{G}/_{A}$ $^{A}/_{C}$ CC 3′ motif. LF-B1 binds to the palindrome 5′ TGGTTAAT/ATTCACCA 3′ which is present in the human α1-antitrypsin gene between positions −78 and −62 from the start of transcription. LF-B1 also recognizes a related sequence present in the human albumin gene between −66 and −50. These results suggest that LF-A1 and LF-B1 are common positive *trans*-acting factors which are required for the expression of several genes in the hepatocyte.

Key words: albumin/α1-antitrypsin/DNA-binding proteins/liver-specific genes/transcriptional control

Introduction

In a multicellular organism many genes are under tight developmental control and are transcribed only in specific, differentiated cell types. The molecular mechanisms by which tissue-specific transcriptional control is achieved are largely unknown. However, this control is likely to depend, at least to a large extent, on the interaction of regulatory proteins (*trans*-acting factors) with specific DNA sequences (*cis*-acting elements) usually, but not always, located in the 5′-flanking region of the gene (reviewed in Dynan, 1987; Maniatis *et al.*, 1987). Thus far, very few tissue-specific DNA elements have been identified. The hepatocyte-specific expression of the genes encoding the plasma proteins provides a convenient system to study tissue specificity. Using 5′ and 3′ deletions, the shortest segment of the 5′-flanking region which is required for hepatocyte specificity has been defined for several genes. For the albumin gene (Ott *et al.*, 1984; Gorski *et al.*, 1986; Babiss *et al.*, 1987), the retinol-binding protein gene (D'Onofrio *et al.*, 1985) and the β-fibrinogen gene (Courtois *et al.*, 1987) a short fragment (~200 bp) of the regulatory region including the homologous start of transcription is sufficient for cell-type specific expression. For the haptoglobin (Oliviero *et al.*, 1987), apolipoprotein A1 (M.Colombo and R.Cortese, in preparation) and α1-antitrypsin (α1AT) genes (De Simone *et al.*, 1987) the region surrounding the TATA box is not required and hepatocyte-specific expression can be obtained with a segment of the 5′-flanking region fused to a heterologous promoter. The DNA−protein interactions within the regulatory region are complex. For instance, the first 200 bp of the rat albumin promoter contain at least six binding sites for factors which are present in rat liver nuclear extracts (Babiss *et al.*, 1987; Cereghini *et al.*, 1987; Lichtsteiner *et al.*, 1987). We have studied the human α1AT gene and defined several DNA elements which are required for efficient transcription in hepatocytes (Ciliberto *et al.*, 1985; De Simone *et al.*, 1987).

α1AT is one of the major protease inhibitors in the blood (for a review see Laurell and Jeppson, 1975). α1AT is expressed in hepatocytes and to a lesser extent in macrophages (Perlmutter *et al.*, 1985). In these two cell types the same coding region is transcribed from two different promoters which are 2 kb apart (Perlino *et al.*, 1987). We have shown that a DNA segment of the hepatocyte specific promoter region between nucleotides −137 and −37 from the start of transcription is sufficient to drive hepatocyte-specific expression not only from the homologous α1AT promoter but also from the heterologous SV40 promoter (De Simone *et al.*, 1987). Within this segment there are two functional domains as defined by mutation analysis, the A-domain from −125 to −100 and the B-domain from −80 to −60 which are both required for expression in hepatoma cells and both of which bind factors present in liver nuclear extracts (De Simone *et al.*, 1987).

Here we report the characterization of two factors, LF-A1 and LF-B1, which bind to the A- and B-domains of the human α_1AT gene respectively. We establish a correlation between binding of these factors and transcriptional activity. We show that LF-A1 also interacts with the regulatory regions of the human apolipoprotein A1 and haptoglobin-related genes, while LF-B1 binds to the human albumin promoter region.

Results

Partial purification of LF-A1 and LF-B1

Previous studies have defined the domains which play an essential role in the expression of the α1AT in hepatoma cells by site-directed mutagenesis (De Simone *et al.*, 1987). The mutants which have been constructed are shown in Figure 1. The EM-3, EM-4 and PM-1 mutations almost completely abolish transcription from the α1AT promoter *in vivo* (<5% of the wild-type activity). The mutants EM-2, EM-5 and PM-2 are reduced in activity (30−50%), while

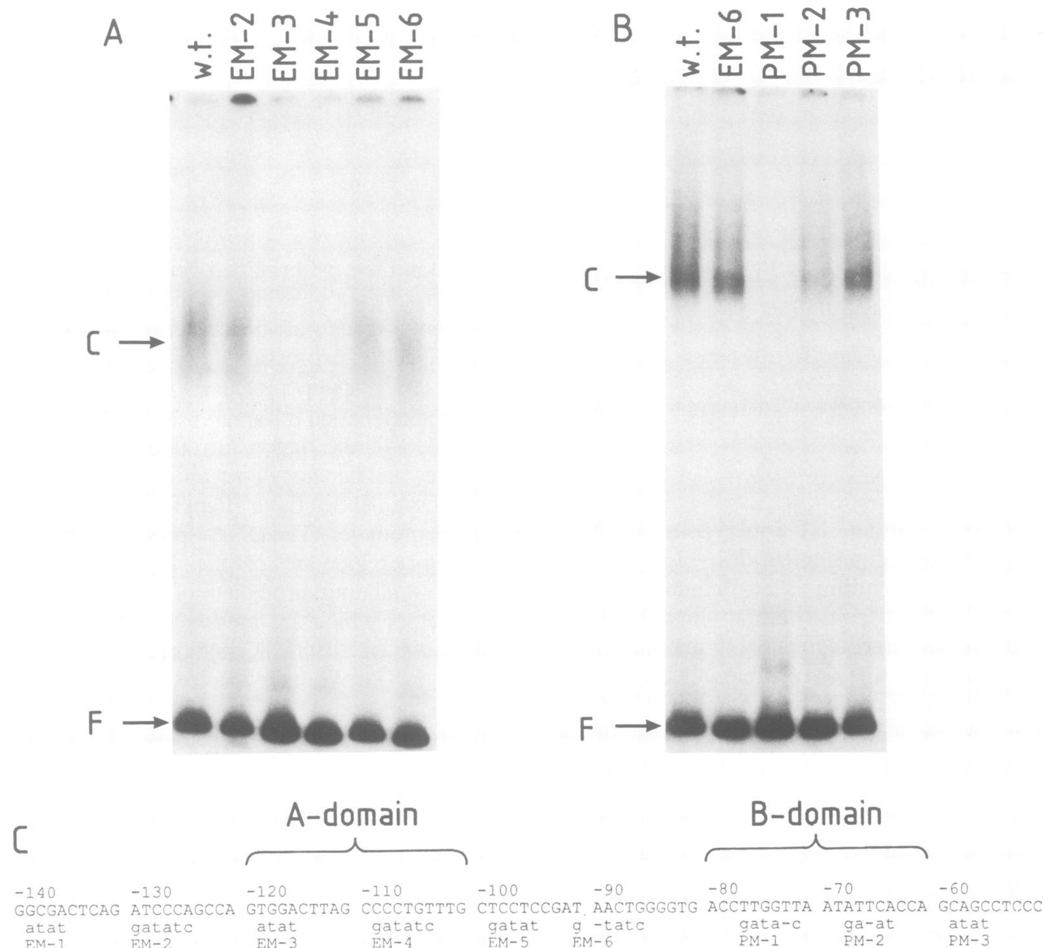

Fig.1. Binding of LF-A1 and LF-B1 correlates with activity *in vivo*. (**A**) Schematic representation of the nucleotides substituted in the EM and PM mutants (De Simone *et al.*, 1987). (**B**) and (**C**) Binding of LF-A1 and LF-B1 to a −261/−37 fragment containing the mutations described in (**A**). F indicates free DNA, C, DNA−protein complexes.

the EM-1, EM-6 and PM-3 mutants display almost wild-type activity (75−80%). Since the EM-3, EM-4 and PM-1 mutations are separated by the EM-5 and EM-6 mutations whose transcriptional activity is affected much less we assume that they represent separate domains which we refer to as the A- and B-domains respectively.

To facilitate the identification of factors which bind to the A- and B-domains of α1AT we synthesized double-stranded (ds) oligonucleotides corresponding to bases −96 to −132 (A-oligo) and −92 to −64 (B-oligo) from the start of transcription. DNA-binding activities were monitored by gel retardation assays (Fried and Crothers, 1981; Schneider *et al.*, 1986).

We have used rat liver as a source of nuclear extract to identify DNA-binding proteins which interact with these domains. The use of heterologous material is justified by the observation that *in vitro* transcription of constructs containing the α1AT promoter in rat nuclear extracts yields results which are comparable with those obtained *in vivo* after transfection of human hepatoma cells (Monaci *et al.*, 1988). Nuclear extracts were prepared according to Dignam *et al.* (1983). The fractionation is described in detail in Materials and methods. The purification procedure can be summarized as follows: the crude nuclear extract was passed over a DEAE−Sepharose column to remove contaminating nucleic acids. No detectable DNA-binding activity remained on the column under the conditions used. The flow-through from the DEAE−Sepharose was fractionated over heparin−Sepharose (HepS). The major activity binding to the A-oligo (called LF-A1) eluted at 600 mM KCl. The LF-A1 material yields a retarded band with a rather smeary appearance, probably due to a poor resolution of multiple forms of the factor irrespective of whether fresh or frozen liver is used. These forms may be either modifications or proteolytic degradation products. However, since the different forms have identical binding properties and co-purify we refer to the entire complex as LF-A1. The major activity which binds to the B-oligo (called LF-B1) eluted at 350 mM KCl (data not shown). The active fractions of each factor were pooled and tested for binding to the mutants in order to establish whether binding of the factor correlated with transcriptional activity. For some experiments the HepS pool of LF-A1 was further purified on an oligonucleotide affinity resin (Kadonaga and Tjian, 1986). The HepS material of LF-B1 was further purified on an FPLC MonoQ column.

Characterization of the binding to the α1AT A- and B-domain

The activity which binds to the A-oligo was tested for binding to a wild-type α1AT promoter fragment comprising nucleotides −264 to −37 and to fragments containing the

Fig. 2. DNase I protection and DMS methylation interference pattern of LF-A1 on α1AT. (**A**) Footprints were performed in the absence (−) or presence (+) of 0.2 μg affinity purified LF-A1. The arrow indicates a DNase I hypersensitivity site, square brackets indicate the extent of DNase I protection. S is a G+A sequence lane. (**B**) Methylation interference pattern of the binding of LF-A1. **Lane F** contains free DNA, **lane C** contains DNA recovered from the DNA−protein complex. The guanine residues whose methylation interferes with binding of LF-A1 are indicated. (**C**) Schematic representation of the data in (**A**) and (**B**). Brackets indicate the extent of DNase I protection, the arrow a DNase I hypersensitivity site, asterisks the guanine residues whose methylation interferes with binding.

EM-2 to EM-6 mutations (Figure 1A). The EM-3 and EM-4 mutations which completely abolish transcriptional activity in hepatoma cells also abolish binding of the activity present in the HepS fraction. This is an indication that the binding activity which we have identified may represent a positive factor whose binding is essential for transcription of α1AT in liver. We have called this factor LF-A1 (liver factor A1). However, the EM-2 and EM-5 mutations which mildly affect transcription *in vivo* do not seem to affect the binding of LF-A1.

In order to define the exact binding site of LF-A1 on α1AT we performed DNase I footprinting (Galas and Schmitz, 1978) and methylation interference (Siebenlist and Gilbert, 1980) experiments using affinity purified material. Figure 2A shows that the region which is protected against DNase I digestion extends from −128 to −107 on the coding (bottom) strand and from −125 and −103 on the non-coding (top) strand, with a DNase I hypersensitivity at −127. Within this region methylation of the guanine residues at position −117 of the coding strand and −120, −119 and −113 of the non-coding strand abolishes the binding while methylation of the G residues at −112, −111 and −110 of the coding strand decreases the binding (Figure 2B). A schematic representation of the DNase I and methylation interference data is given in Figure 2C.

The activity which binds to the B-oligo was tested for binding to the wild-type fragment and to the mutants EM-6, and PM-1 to PM-3. The mutant PM-1 which is no longer transcribed *in vivo* is unable to bind this activity *in vitro* (Figure 1B). Similarly, the mutant PM-2 which is less active in transcription binds less efficiently. The flanking mutants EM-6 and PM-3 whose transcriptional activity is unaffected bind LF-B1 like the wild-type promoter fragment. We have called the factor LF-B1, and believe that like LF-A1, LF-B1 is a positive regulator of α1AT. The DNase I protection of LF-B1 on α1AT extends from −86 to −59 on the coding strand and from −83 to −56 on the non-coding strand, with a weak hypersensitivity site at −84 (Figure 3A). Within this region there are several G residues whose methylation interferes with the binding, namely the guanines at positions −67, −65 and −64 on the coding strand and at −77 and −76 on the non-coding strand (Figure 3B). The LF-B1 binding site contains a near perfect (7/8) palindrome 5′ TGGTTAAT/ATTCACCA 3′ centred between nucleotides −71/−70 (Figure 3C). The entire palindrome is involved in the binding of LF-B1 since G residues in both halves are contacted by the factor.

A useful property of LF-B1 is the relative thermostability

Fig. 3. DNase I protection and DMS methylation interference pattern of LF-B1 on α1AT. (**A**) Footprint reactions were performed in the absence (−) or presence (+) of partially purified LF-B1 (2 μg of a 350 mM HepS pool). The arrow indicates a DNase I hypersensitivity site, brackets indicate the extent of DNase I protection. S is a G+A sequence lane. (**B**) Methylation interference pattern of the binding of LF-B1 on α1AT. **Lane F** contains free DNA, **lane C** contains DNA recovered from the DNA−protein complex. Guanine residues whose methylation interferes with binding of LF-B1 are indicated. (**C**) Schematic representation of the data in (**A**) and (**B**). Brackets indicate the extent of DNase I protection, the vertical arrow a DNase I hypersensitivity site, asterisks the guanine residues whose methylation interferes with binding. The palindrome is denoted by horizontal arrows.

of its binding activity. Heating for 5 min at 60°C which completely abolishes binding of LF-A1 (Figure 4A) does not affect the binding of LF-B1 (Figure 4B). Only at higher temperatures (e.g. 90°C) is the binding activity reduced and lower bands appear, presumably caused by degradation products of LF-B1.

Binding of LF-A1 and LF-B1 to the regulatory regions of other human liver-specific genes

To determine whether the binding of LF-A1 and LF-B1 is specific for α1AT we tested these factors for binding to the 5′-flanking regions of the apolipoprotein A1 (ApoA1; −255/−6), retinol binding protein (RBP; −322/−7), haptoglobin-related (Hpr; −183/+44) and albumin (Alb; −221/−1) genes using gel retention as an assay. These fragments contain the minimal segment required for hepatocyte-specific transcription of the genes. With the partially purified LF-A1 material we noticed a binding to the ApoA1 and Hpr fragments while the LF-B1 material bound to the Alb fragment. To ascertain that the band-shifts which we obtained are due to the specific binding of LF-A1 or LF-B1 and not to any non-specific interaction or contaminating activities we performed the competition experiments shown in Figure 5. The protein fractions were pre-incubated with a 100-fold molar excess of unabelled specific oligo (A-oligo for LF-A1 and B-oligo for LF-B1) or a non-specific ds oligo before addition of labelled fragment (see Materials and methods). Binding to the α1AT fragment and to both the ApoA1 and Hpr fragments can be competed with the A-oligo but not with the same amount of non-specific oligo. Since LF-A1 is the only protein shown to bind the A-oligo in the material used, these results strongly suggest that LF-A1 binds to all three genes. The amount of DNA retarded with the ApoA1 probe is reproducibly greater

than with the α1AT probe, while Hpr generates a weaker shift. LF-A1 also binds, although less efficiently, to the haptoglobin (Hp1) 5'-flanking region which is highly homologous to Hpr (data not shown). If we presume that the intensity of the retarded band reflects the affinity of the protein for a sequence the relative affinities of LF-A1 for these sites are: ApoA1 > α1AT > Hpr > Hp1.

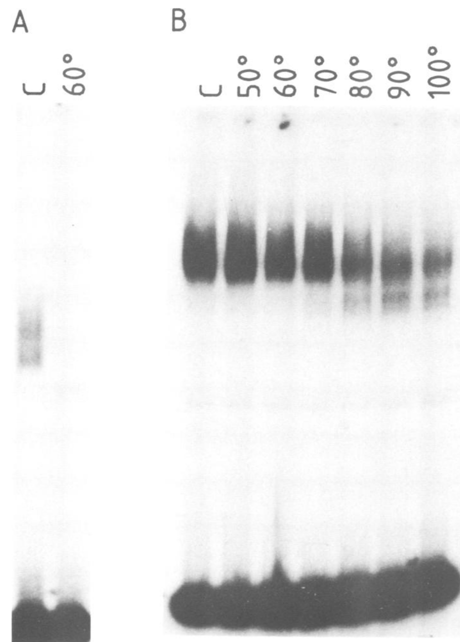

Fig. 4. Heat resistance of the LF-A1 and LF-B1 binding activities. Aliquots of a HepS fraction of LF-A1 (**panel A**) and a MonoQ fraction of LF-B1 (**panel B**) were heated for 5 min at the temperatures indicated. The precipitate of denatured protein was pelleted and the supernatant assayed for binding to a wild-type −261/−37 fragment of the α1AT promoter region. **Lane C** contains a non-heated control.

Similarly, the binding of LF-B1 to the α1AT and albumin fragment can be competed with the B-oligo but not with the same quantity of non-related oligo. Furthermore the heat inactivation profiles of the binding to α1AT and albumin are identical (data not shown). Hence we conclude that LF-B1 binds to the albumin promoter region.

We determined the precise binding site of LF-A1 to ApoA1 and LF-B1 to albumin by DNase I protection and methylation interference experiments (Figures 6 and 7). The protection from DNase I digestion of LF-A1 on ApoA1 extends from −221 to −198 on the coding strand and from −215 to −195 on the non-coding strand, with a weak hypersensitivity site at −216 (Figure 6A). Within this region methylation of the guanine residues at −210, −209 and −203 to −201 on the coding strand and −213 and −205 on the non-coding strand interferes with the binding of LF-A1 (Figure 6B). These results are summarized in Figure 6C. Comparison of the binding sites of LF-A1 on α1AT and ApoA1 reveals a common TG$^{G}/_{A}$ $^{C}/_{A}$ CC motif which is present as a tandem repeat in both sites. Methylation of guanine residues in any of these motifs interferes with binding. The DNase I protection of LF-B1 on the albumin gene extends from −72 to −46 on the coding strand and from −69 to −42 on the non-coding strand (Figure 7A). Methylation of the guanine residues at the positions −54 and −51 on the coding strand or −63 on the non-coding strand abolishes binding of LF-B1 (Figure 7B). This shows

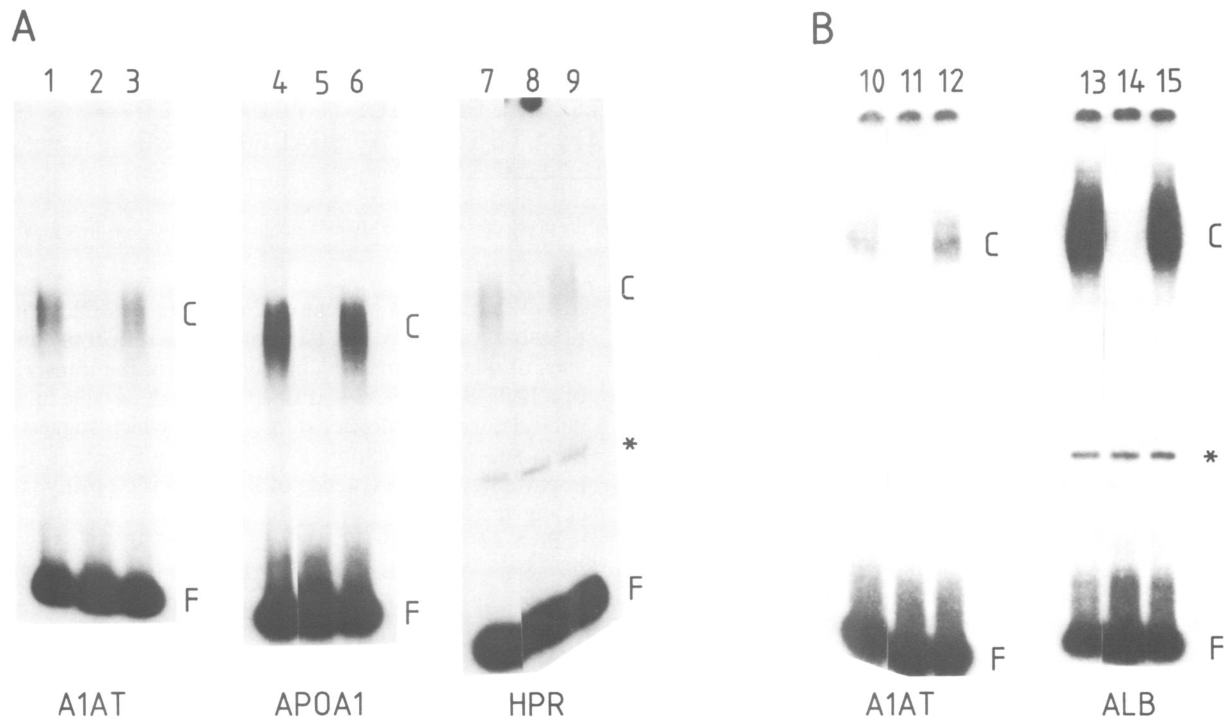

Fig. 5. Binding of LF-A1 and LF-B1 to the regulatory regions of other genes. 0.02 μg of affinity purified LF-A1 (**panel A**) or 0.4 μg of MonoQ purified LF-B1 (**panel B**) were pre-incubated without competitor (**lanes 1, 4, 7, 10** and **13**), with specific competitor (**lanes 2, 5, 8, 11** and **14**) or with non-specific competitor (**lanes 3, 6, 9, 12** and **15**). For LF-A1 the A-oligo was used as a specific competitor and for LF-B1 the B-oligo. A ds oligonucleotide of similar length was used as non-specific competitor (see Materials and methods). F indicates free DNA, C DNA−protein complex. The asterisk indicates an extra band which is due to partial denaturation of the probe during ethanol precipitation (Svaren *et al.*, 1987). The presence of this band does not affect the binding of LF-A1 or LF-B1.

Fig. 6. DNase I protection and methylation interference pattern of LF-A1 on the apolipoprotein A1 5′-flanking region. Reaction conditions and use of symbols as in Figure 2.

that LF-B1 binds to a region which contains the sequence 5′ TAGTTAATAATCTACA 3′. The 5′ part of this sequence strongly resembles the α1AT half sites (TGGT^{T}/$_{G}$ AAT).

Discussion

We have previously shown that a segment from the human α1AT promoter comprising nucleotides −137 to −37 from the start of transcription is able not only to direct tissue-specific expression from the homologous α1AT promoter but also to confer tissue-specificity on the heterologous SV40 promoter (De Simone *et al.*, 1987). Within this fragment we have defined two regions (referred to as the A- and B-regions) which are essential for expression in hepatoma cells. Here we have characterized two proteins (LF-A1 and LF-B1) isolated from rat liver extracts which bind to these regions. For LF-B1 we have established a good correlation between the ability of mutated promoter regions to bind the factor and transcriptional activity of these mutants in hepatoma cells. In the case of LF-A1, however, mutations outside the region which we have shown to be important for factor binding also affect transcription *in vivo* although considerably less. We cannot at present exclude other possible explanations but we believe that these mutations reflect the binding sites of other factors which play a role in the transcription of α1AT. Both LF-A1 and LF-B1 are undetectable in spleen nuclear extracts (Monaci *et al.*, 1988), which suggests that the tissue-specific expression of the α1AT gene is at least in part due to the existence of liver-specific positive *trans*-acting factors.

Both LF-A1 and LF-B1 bind to the 5′-flanking regions of genes other than α1AT; LF-A1 interacts with the 5′-flanking region of the ApoA1 and Hpr genes while LF-B1 binds to the albumin regulatory region. The criteria which we applied to decide that the same factor binds to two or more genes were: (i) similarity of the band shifts (complexes migrate at similar positions); (ii) co-purification of the binding activities; (iii) competition of the binding to one gene with a synthetic ds oligo which carries the binding site but not with a non-specific ds oligo; (iv) homology of the binding sites; (v) similarity of the footprint and methylation interference patterns; (vi) identical heat-resistance of the binding (for LF-B1 only).

Recently, Courtois *et al.* (1987) reported the identifica-

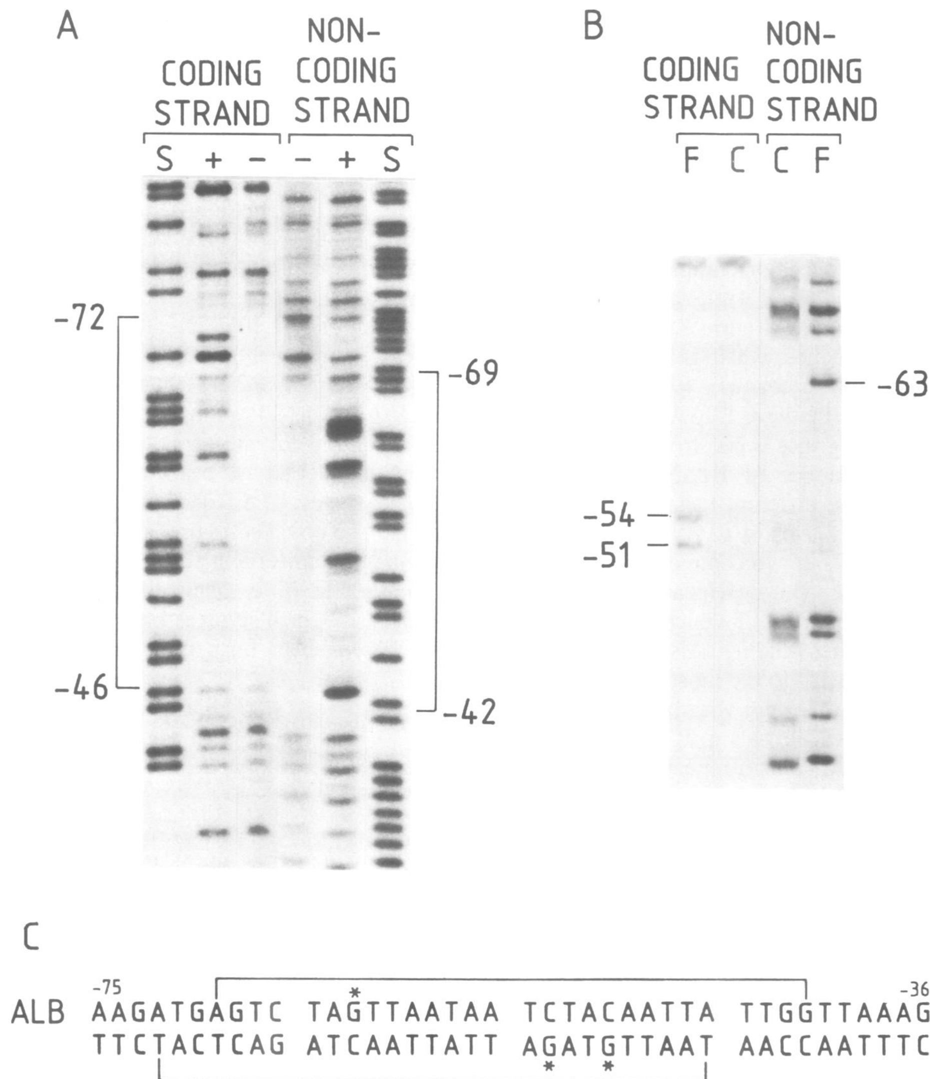

Fig. 7. DNase I protection and methylation interference pattern of LF-B1 on the albumin 5′-flanking region. Reaction conditions and use of symbols as in Figure 3.

tion of a protein (HNF1) from rat hepatocyte extract which binds to the α1AT B-region and also to the rat α- and β-fibrinogen promoter region. These authors propose the sequence ATTAAC (or GTTAAT) as a consensus for the binding site of HNF1. This hexamer is part of the larger consensus sequence which we derived from the binding sites of LF-B1. We tested a synthetic ds oligo corresponding to the binding site of HNF1 on β-fibrinogen (corresponding to the nucleotides −95 to −76 from the start of transcription) and observed a strong binding of LF-B1 to this sequence (E.M.Hardon, unpublished observations). From these data it seems that LF-B1 and HNF-1 are identical. Lichtsteiner *et al.* (1987) have identified a protein (which they call B-protein) which binds to the mouse albumin promoter region and generates a DNase I footprint which is very similar to the footprint obtained from LF-B1 on the the human gene. We therefore believe LF-B1, HNF1 and the B-protein to be identical. However, since formal proof is lacking we prefer to call the factor which we have identified LF-B1.

Comparison of the binding sites of LF-B1 on the human albumin and α1AT genes with the homologous sites on the same genes from other species and the α- and β-fibrinogen sites shows a sequence 5′ T $^G/_A$ GTTAAT 3′ which is

A1AT	human	-80	CTTGGTTAAT	ATTCACCAGC	-61
	mouse		ATTGGTTAAT	ATTCATAGCA	
ALB	human	-67	TCTAGTTAAT	AATCTACAAT	-48
	mouse		TATGGTTAAT	GATCTACAGT	
	rat		TGTGGTTAAT	GATCTACAGT	
A-FIBR	rat	-42	CTAGGTTAAT	CATCACCCTT	-61
B-FIBR	rat	-79	TTTAGTTAAT	ATTTGACAGT	-98
CONS			TGGTTAAT	NTTCNNCA A	

Fig. 8. Comparison of LF-B1 binding sites.

strongly conserved. This sequence is followed by a degenerate version of the same motif (TGN $^T/_A$ G<u>A</u>$^A/_T$N) in the opposite orientiation. In the second (3′) half site only the A residue which is underlined is fully conserved, while both G residues are strongly conserved (6/7). The entire palindrome is involved in binding as judged from the methylation interference data. However, both the higher degree of conservation and the observation that the PM-2 mutation in the second (3′) half of the α1AT LF-B1 site is less severe than the PM-1 mutation in the 5′ half site, suggest that the 5′ half site is more important for binding.

The LF-B1 site shows some similarity to the binding site of other transcription factors, most noticeably the core enhancer motif (GTGG $^{A}/_{T}{}^{A}/_{T}{}^{A}/_{T}$G) (Weiher *et al.*, 1982) and the NF-Y recognition sequence (CTGATTGG $^{C}/_{T}{}^{C}/_{T}$) (Dorn *et al.*, 1987; Hooft *et al.*, 1987). However, on the basis of careful sequence comparison and mutational analysis we can exclude the possibility that LF-B1 is identical to either the core binding protein or NF-Y. Mutation of the first 4 bp of the NF-Y site strongly reduces or completely abolishes binding of NF-Y (Dorn *et al.*, 1987), while these nucleotides are not conserved in LF-B1 sites (Figure 8). Likewise the first G residue of the core enhancer which is fully conserved in the known binding sites (Barrett *et al.*, 1987) is not conserved in the LF-B1 sites while the GTTAAT motif which is present in all LF-B1 binding sites is not strongly conserved in the core enhancer sites. Hence LF-B1, NF-Y and core enhancer binding protein are different factors.

Comparison of the binding sites of LF-A1 in the α1AT and ApoA1 promoter regions shows a common 5′ CCCCTG 3′ motif which is also present in the Hp and Hpr genes. However, the EM-3 mutation in α1AT which does not alter this motif completely abolishes binding. Also, methylation of the G residue complementary to the most 3′ C or of the G within this motif does not interfere with binding. Closer inspection of the binding sites reveals a 5′ TG $^{G}/_{A}{}^{A}/_{C}$CC 3′ motif which is present as a tandem repeat in both genes although with different spacing. Methylation of any G residue within this motif does interfere with binding. The same motif is present in a single copy of the Hpr regulatory region around position −90 and again around −140. We have indications that LF-A1 binds to both Hpr domains, although with a low affinity (E.M.Hardon, unpublished observations). This could be explained by the occurrence of a single copy of the motif rather than a tandem repeat. The regulatory region of the haptoglobin gene (Hp1) is highly homologous to the regulatory region of the Hpr gene [only 8 bp are different in the first 184 bp from the start of transcription (Oliviero *et al.*, 1987)]. One of these substitutions (a T for a C at position 137) changes the distal copy of the TG $^{G}/_{A}{}^{A}/_{C}$ CC motif, while the proximal copy remains intact. As would have been predicted if this motif plays a role in the binding, Hp1 binds LF-A1 less strongly than Hpr. The weaker binding correlates with lower expression *in vivo*. The binding sites of LF-A1 do not show an obvious sequence homology with the sites of known transcription factors.

The liver-specific genes characterized thus far differ with respect to the contribution of the DNA sequences surrounding the TATA box to tissue-specific expression. The upstream sequences of the albumin (Gorski *et al.*, 1986; Frain *et al.*, in preparation) and retinol binding protein genes (D'Onofrio *et al.*, 1985; Colantuoni *et al.*, 1987) are expressed in a tissue-specific manner only with their own promoter or another liver-specific promoter. In contrast, ApoA1 (M.Colombo and R.Cortese, in preparation) and Hpr (Oliviero *et al.*, 1987) behave like α1AT in that their upstream sequences activate the SV40 promoter in hepatoma cell lines. Interestingly, we have detected binding of LF-A1 only with those 5′-flanking regions which are able to confer liver specificity on a heterologous promoter. The presence of a binding site for LF-A1 is not sufficient for expression in hepatoma cell lines since the ability to bind both LF-A1 and LF-B1 is required for tissue-specific transcription of α1AT *in vivo* (De Simone *et al.*, 1987).

α1AT is, however, the only gene identified thus far to contain binding sites for both LF-A1 and LF-B1. It has been shown that at least two factors bind to the shortest fragment of the ApoA1 gene which is able to confer tissue-specificity on a heterologous promoter (Monaci *et al.*, unpublished observations). We have identified one of these factors at LF-A1, but our binding data demonstrate that the other factor is not LF-B1. The regulation of the albumin gene seems to be more complex in that many factors have been shown to bind to the promoter region (Babiss *et al.*, 1987; Cereghini *et al.*, 1987; Lichtsteiner *et al.*, 1987). Nevertheless we find only a single binding site for LF-B1 in the proximal region (−221/+1) and we do not observe binding of LF-A1. This shows that although LF-A1 and LF-B1 are both required for the expression of the α1AT gene, each factor can interact independently with the regulatory regions of other genes. The requirement for either of these factors can thus be obviated in the presence of other factors which may be able to serve a similar function. This suggests that liver specificity is obtained through a highly modular mechanism which involves factors binding to the TATA box and to sequences further upstream.

Materials and methods

Nuclear extract preparation and protein purification

Nuclear extracts were prepared from frozen rat liver essentially as described by Dignam *et al.* (1983). The crude nuclear extract was dialysed extensively against buffer D [20 mM Hepes pH 7.9, 10% glycerol, 0.2 mM EDTA, 0.5 mM (DTT) and 1 mM phenylmethylsulphonyl fluoride] with 300 mM KCl, and centrifuged for 10 min at 12 000 r.p.m. in a Sorvall SS34 rotor to remove the precipitate. The clear supernatant was passed through DEAE−Sepharose and the flow-through was diluted with 0.5 vol buffer D and loaded onto a heparin−Sepharose column. The HepS column was eluted with a linear gradient of KCl in buffer D. The pooled active fractions of LF-A1 which elutes at 600 mM KCl were dialysed against buffer D with 100 mM KCl and loaded onto an oligo-affinity column (Kadonaga and Tjian, 1986) which contained 4−8 repeats of the sequence 5′ CGCCCCCACT GAACCCTTGA CCCCTGCCCT CGCC 3′, corresponding to the LF-A1 site of the ApoA1 gene. This column was eluted with buffer D containing 500 mM KCl and the active fractions were pooled. The HepS fractions containing the LF-B1 activity which elutes at 350 mM KCl were dialysed against buffer D with 100 mM KCl, loaded on an FPLC MonoQ column and eluted with a linear gradient of KCl in buffer D. The active fractions of LF-B1 which elutes from MonoQ at 270−280 mM KCl were pooled. For some experiments the protein material was concentrated using centricon (Amicon) cartridges.

End-labelled DNA probes

Fragments for end-labelled probes were excised from pEMBL-CAT (M.Uhlen, in preparation; Colantuoni *et al.*, 1987) constructs containing promoter segments comprising the nucleotides indicated between brackets. At the 5′ side of the fragment the polylinker sites of pEMBL-CAT were used and at the 3′ side either the *Hin*dIII site flanking the CAT gene or, in the case of α1AT, the naturally occurring *Bam*HI site.

The α1AT *Kpn*I/*Bam*HI promoter fragment (−264/−37) (De Simone *et al.*, 1987) was end-labelled at the *Bam*HI site either with Klenow enzyme and [α-^{32}P]dATP or kinase and [γ-^{32}P]ATP.

The ApoA1 *Xma*I/*Hin*dIII fragment (−255/−6) (Colombo and Cortese, in preparation) was labelled at the *Xma*I site. Both the albumin *Dra*I/*Hin*dIII fragment (−221/−1) (Urano *et al.*, 1986) and the Hpr *Sma*I/*Hin*dIII fragment (−183/+44) (Oliviero *et al.*, 1987) were labelled at the *Hin*dIII site.

Gel retention assays

For gel retention assays partially purified proteins were pre-incubated in a 20 μl reaction containing 20 mM Tris pH 7.6, 8% Ficoll, 50 mM KCl, 1 mM EDTA, 0.2 mM DTT, 100 ng sonicated salmon sperm DNA, 3 μg poly(dIdC) and for LF-B1 5 mM spermidine. After 10 min 10 000−20 000 c.p.m. (Cerenkov) end-labelled DNA fragment was added and the incubation was continued for 10 min at room temperature. Free DNA and DNA−protein complexes were resolved on a 5% polyacrylamide gel in

0.5 × TBE (45 mM Tris-borate, 45 mM boric acid and 2 mM EDTA). For competition experiments the competitor was included during the pre-incubation. The ds oligonucleotide 5′ GCGGGGAGAC CTAGGTGAC GAATTCCTAG GGCCC 3′ was used as a non-specific competitor.

DNase I footprint assays

Protein fractions were pre-incubated in a 10 μl reaction containing 20 mM Hepes pH 7.9, 50 mM KCl, 0.2 mM EDTA, 10 ng pUC8 and for LF-B1 3 μg poly(dIdC). After 10 min on ice 5000 – 10 000 c.p.m. (Cerenkov) end-labelled probe was added and the incubation continued for 10 min at room temperature. 2 μl of DNase I, freshly diluted to a final concentration of 0.2 – 20 μg/ml in 10 mM $CaCl_2$, was added and the digestion was allowed to proceed for 90 s at room temperature. Digestion was stopped by the addition of 50 μl phenol – chloroform and 40 μl 0.3 mM NaAc and 1% SDS. The DNA was extracted once with phenol/chloroform, precipitated with 2.5 vol EtOH, resuspended in 80% formamide and electrophoresed on a 6% acrylamide/7 M urea gel.

Methylation interference assays

Protein fractions were incubated with DNA under similar conditions as for gel retention assays with the difference that a larger amount of partially methylated (Maxam and Gilbert, 1980) probe (60 000 – 100 000 c.p.m.) was added. Both the free DNA and the DNA – protein complex were excised from the gel and the DNA was eluted in 0.5 × Maxam and Gilbert buffer (Maxam and Gilbert, 1980). The DNA was purified over elutip mini columns (Schleicher and Schuell), precipitated with EtOH, treated with 10% piperidine for 30 min at 90°C, dried and electrophoresed on 6% acrylamide/7 M urea gels. Quantitative densitometric scanning was used to analyse the autoradiogram.

Acknowledgements

We wish to thank Carolyn Blance for excellent technical assistance and Graham Tebb, Henk Stunnenberg and Paolo Di Nocera for critically reading the manuscript and many other members of the Gene Structure and Regulation Programme for their helpful comments and suggestions. Part of this work was supported by EEC grant BAP-0115-D.

References

Babiss,L.E., Herbst,R.S., Bennett,A.L. and Darnell,J.E. (1987) *Genes Dev.*, **1**, 256 – 267.

Barrett,P., Clark,L. and Hay,R.T. (1987) *Nucleic Acids Res.*, **15**, 2719 – 2734.

Cereghini,S., Raymondjean,M., Carranca,A.J., Herbomel,P. and Yaniv,M. (1987) *Cell*, **50**, 627 – 638.

Ciliberto,G., Dente,L. and Cortese,R. (1985) *Cell*, **41**, 531 – 540.

Colantuoni,V., Pitozzi,A. Blance,C. and Cortese,R. (1983) *EMBO J.*, **2**, 631 – 636.

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

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

D'Onofrio,C., Colantuoni,V. and Cortese,R. (1985) *EMBO J.*, **4**, 1981 – 1989.

Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1476 – 1489.

Dorn,A., Bollekens,J., Staub,A., Benoist,C. and Mathis,D. (1987) *Cell*, **50**, 863 – 872.

Dynan,W.S. (1987) *Trends Genet.*, **3**, 121 – 122.

Fried,M. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505 – 6525.

Galas,D. and Schmitz,A. (1978) *Nucleic Acids Res.*, **5**, 3157 – 3170.

Gorski,K., Carneiro,M. and Schibler,U. (1986) *Cell*, **47**, 767 – 776.

Hooft,R.A.M., Bollekens,J., Dorn,A., Benoist,C. and Mathis,D. (1987) *Nucleic Acids Res.*, **15**, 7265 – 7281.

Kadonaga,J.T. and Tjian,R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5889 – 5893.

Laurell,C.B. and Jeppson,J.O. (1975) In Putman,F.W. (ed.), *The Plasma Proteins*. 2nd edn, Academic Press, NY, Vol II, pp. 229 – 264.

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

Maniatis,T., Goodburn,S. and Fischer,J.A. (1987) *Science*, **236**, 1237 – 1244.

Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499 – 560.

Monaci,P., Nicosia,A. and Cortese,R. (1988) *EMBO J.*, **7**, in press.

Oliviero,S., Morrone,G. and Cortese,R. (1987) *EMBO J.*, **6**, 1905 – 1912.

Ott,M., Sperling,L., Herbomel,P., Yaniv,M. and Weiss,M.C. (1984) *EMBO J.*, **3**, 2505 – 2510.

Perlino,E., Cortese,R. and Ciliberto,G. (1987) *EMBO J.*, **6**, 2767 – 2771.

Perlmutter,O.H., Kay,R.M., Cole,S.F., Rossin,T.H., Van Thiel,D. and Colten,H.R. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6918 – 6921.

Schneider,R., Gander,I., Müller,U., Mertz,R. and Winnacker,E.L. (1986) *Nucleic Acids Res.*, **14**, 1303 – 1317.

Svaren,J., Inigami,S., Lovegren,E. and Chalkley,R. (1987) *Nucleic Acids Res.*, **15**, 8739 – 8752.

Siebenlist,U. and Gilbert,W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 122 – 126.

Urano,Y., Watanabe,K., Sakai,M. and Tamaoki,T. (1986) *J. Biol. Chem.*, **261**, 3244 – 3251.

Weiher,H., Konig,M. and Gruss,P. (1982) *Science*, **219**, 626 – 631.

Received on February 5, 1988; revised on March 21, 1988