MOLECULAR AND CELLULAR BIOLOGY, July 1993, p. 3860–3871
0270-7306/93/073860-12$02.00/0

Vol. 13, No. 7

The Lung-Specific CC10 Gene Is Regulated by Transcription Factors from the AP-1, Octamer, and Hepatocyte Nuclear Factor 3 Families

PATRICIA L. SAWAYA,[1] BARRY R. STRIPP,[2] JEFFREY A. WHITSETT,[2] AND DONAL S. LUSE[1]*

Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati, Cincinnati, Ohio 45267-0524,[1] and Division of Pulmonary Biology, Children's Hospital Research Foundation, Cincinnati, Ohio 45229[2]

Received 17 November 1992/Returned for modification 4 January 1993/Accepted 12 April 1993

We have shown that a large fragment (−2339 to +57) from the rat CC10 gene directed lung-specific expression of a reporter construct in transgenic animals. Upon transfection, a smaller fragment (−165 to +57) supported reporter gene expression exclusively in the Clara cell-like NCI-H441 cell line, suggesting that a Clara cell-specific transcriptional element resided on this fragment (B. R. Stripp, P. L. Sawaya, D. S. Luse, K. A. Wikenheiser, S. E. Wert, J. A. Huffman, D. L. Lattier, G. Singh, S. L. Katyal, and J. A. Whitsett, J. Biol. Chem. 267:14703–14712, 1992). The interactions of nuclear proteins with a particular segment of the CC10 promoter which extends from 79 to 128 bp upstream of the CC10 transcription initiation site (CC10 region I) have now been studied. This sequence can stimulate both in vitro transcription in H441 nuclear extract and transient expression of reporter constructs in H441 cells. Electrophoretic mobility shift assays using extracts from H441, HeLa, rat liver, and fetal sheep lung cells were used to demonstrate that members of the AP-1, octamer, and HNF-3 families bind to CC10 region I. Transcription factors from H441 cells which are capable of binding to CC10 region I are either absent in HeLa, rat liver, and fetal sheep lung extracts or enriched in H441 extracts relative to extracts from non-Clara cells.

Lung tissues contain mRNA for transcription factors previously identified in other organs, including hepatocyte nuclear factor 1β (HNF-1β), HNF-3α and -β, and C/EBPα, -β, and -δ, which are known to be enriched in hepatic and adipose tissues (2, 3, 17, 28, 31, 46). The lung also contains ubiquitous transcription factors, including AP-1 and CREB-2 (21, 46). However, no studies have addressed the localization of these transcription factors to any particular lung cell type. The study of transcriptional regulation in the lung has been complicated by the cellular heterogeneity of this organ. Adult human lungs are composed of over 40 different cell types (11), which has made the preparation of purified epithelial cells for the purpose of biochemical analysis very difficult. Recently, sequences for several lung-specific promoters have been obtained, such as those for the lung surfactant proteins A, B, and C and for CC10, a secretory product of airway cells (for a review, see reference 40). Genetic studies using these sequences have begun to provide information on transcriptional regulation in the lung (12, 22, 25). For example, in situ hybridization in mice and rats has shown that the surfactant protein C gene is expressed exclusively in the alveolar epithelium and the CC10 gene is expressed primarily in the airway, while the surfactant A and B genes are expressed at high levels in both of these locations (12, 20, 45). Defining the temporal and spatial expression patterns of various lung-specific genes in different pulmonary cell types, and determining the basis of their regulation, may help in understanding the developmental and functional relationships between closely related lung cells.

CC10 is an abundant secretory product of Clara cells (36, 37) and is known by several names in the literature, including CC17 and RLL (for rat lung lavage [44]), Clara cell secretory protein (14), and polychlorinated biphenyl-binding protein (30). Clara cells are the primary site of CC10 expression, but CC10 is also produced at a low level in the uterus under some conditions (16). We have previously shown that 2.4 kb of upstream CC10 promoter sequence was sufficient to direct airway-specific expression of a reporter gene in transgenic mice (39). In transient expression assays, 165 bp of upstream promoter sequence directed expression of a reporter gene in NCI-H441 cells, a human lung adenocarcinoma with characteristics of Clara cells, but not in HeLa cells or in A549 cells, a human lung adenocarcinoma which does not resemble Clara cells. DNase I footprinting with H441 nuclear extracts revealed a complex protein binding site, which we termed region I, located 79 to 128 bp upstream of the transcriptional start site. HeLa nuclear proteins also bound this region, but the DNase I protection pattern with HeLa extract was different from that seen with H441 nuclear extract (39).

In this study, in vitro transcription and transient transfection assays were used to functionally characterize regulatory properties of promoter proximal sequences of the rat CC10 gene. Electrophoretic mobility shift assays (EMSAs) utilizing H441, HeLa, rat liver, and fetal sheep lung (FSL) nuclear extracts demonstrated that members of the AP-1, HNF-3, and octamer families are capable of interacting with CC10 region I in these extracts. However, the same members of each factor family did not bind in each extract. In particular, the combination of factors that bound region I in H441 cells did not bind in the non-Clara cells. These data suggest that the cell-specific transcription of the rat CC10 gene in Clara cells could result from the interactions of a specific combination of transcription factors from the AP-1, octamer, and HNF-3 families within region I.

* Corresponding author.

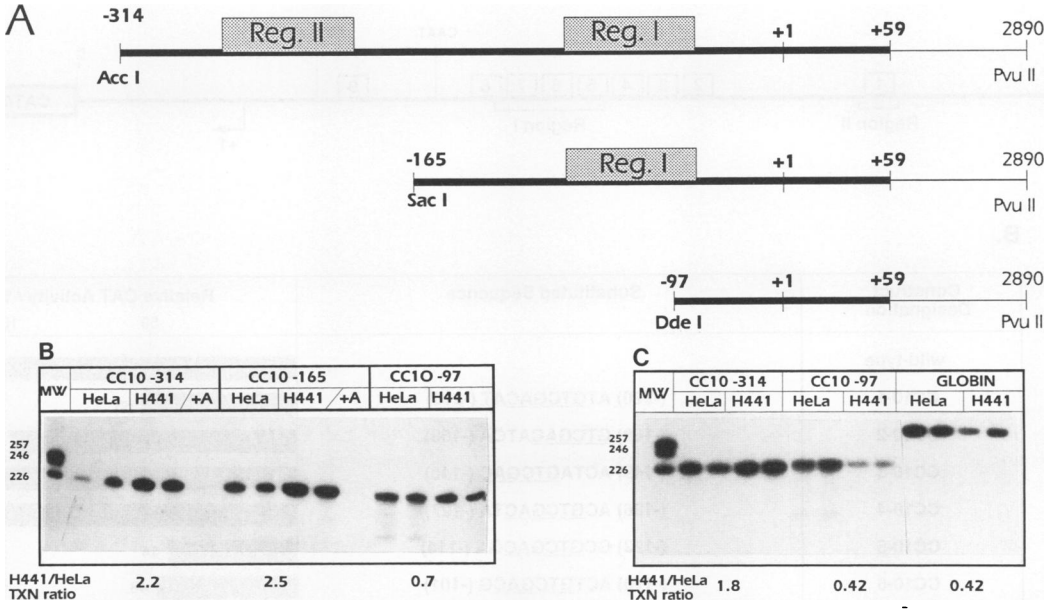


FIG. 1. In vitro transcription, showing that positive transcriptional control elements responsive to factors in H441 extract but not HeLa extract lie upstream of −97 in the rat CC10 promoter. (A) CC10 promoter-bearing templates were prepared from the original CC10 subclone (pGEMRtCC10 [39]) by digestion with the restriction enzymes indicated. The heavy line indicates CC10 sequences; regions which we had previously shown interact with H441 nuclear proteins are designated regions I and II (39). (B and C) Runoff transcripts were generated by using the indicated nuclear extracts as described in detail in Materials and Methods. Size markers are present in lane MW (sizes are indicated in nucleotides); reactions containing α-amanitin at 0.5 μg/ml are marked +A. Note that all of the reactions were done in duplicate except for the α-amanitin controls. Separate preparations of RNA were synthesized with the CC10 −314 promoter and H441 or HeLa extracts for panels B and C. The average ratio of the transcriptional (TXN) activity of the H441 and HeLa extracts on the various promoters is given below each group of lanes; the value for the leftmost −314/HeLa lane in panel B, in which the sample did not efficiently chase (not shown), was not included in the ratio.

MATERIALS AND METHODS

Preparation of nuclear extracts. NCI-H441 and HeLa cell nuclear extracts were prepared as described previously (19, 39). Livers from adult Sprague-Dawley rats were used to prepare nuclear extracts by the method of Gorski et al. (13). FSLs (82 days of gestation; a gift from Harriet Yamamoto) were used to prepare an extract (a gift from Charlene McWhinney) by the method of Tian and Schibler (43).

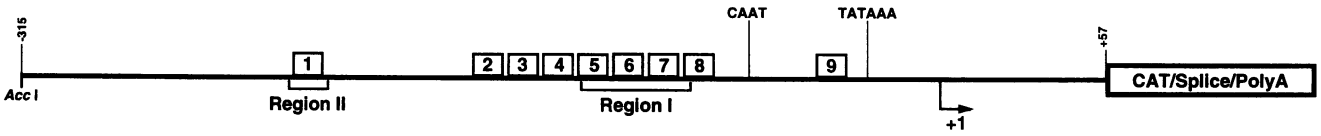
In vitro transcription. CC10 promoter templates for in vitro transcription were made by digesting pGEMRtCC10 (39) with pairs of restriction enzymes to yield promoter-bearing fragments. In all cases, the restriction site downstream of +1 was *Pvu*II (+226); the upstream site was *Acc*I (−314), *Sac*I (−165), or *Dde*I (−97) (Fig. 1A). The minimal globin promoter template was made by digesting pMB5T with *Pvu*II; this promoter contains 124 bp of the mouse beta-globin locus cloned into pUC18 (18). These digested plasmids were used as templates for in vitro transcription without purification of the promoter-bearing fragment; CC10 templates were expected to generate a 226-nucleotide runoff transcript; the globin template was expected to generate a 254-nucleotide transcript.

The total volume of each in vitro transcription reaction was 30 μl. Preinitiation complexes were assembled in 20-μl reaction mixtures containing 16.5 μg of cut plasmid per ml, 5 mM $MgCl_2$, 75 mM KCl, 10% glycerol, and approximately 50 μg of H441 or HeLa nuclear proteins; these reaction mixtures were incubated for 25 min at 30°C. α-Amanitin at 0.5 μg/ml was added to some preinitiation complex assembly reactions to demonstrate that transcription is dependent on RNA polymerase II. Transcription was then performed as a pulse and chase. Pulse reagents were added in 5 μl to bring the reaction mixtures to 1 mM adenylyl cytidine (ApC)–10 μM ATP–0.5 μM UTP–0.25 μM CTP–0.25 μM [α-^{32}P]CTP (800 Ci/mmol). After 5 min at 30°C, 5 μl of chase mix was added to give final concentrations of 0.5% Sarkosyl and 500 μM ATP, CTP, GTP, and UTP. After 10 min at 30°C, the reaction mixtures were brought to 110 μl and digested with 150 μg of proteinase K per ml at 55°C for 20 min. The nucleic acids were then phenol-$CHCl_3$ extracted, ethanol precipitated, and resolved on 10% polyacrylamide sequencing gels containing 8 M urea. The amount of radioactivity in each transcript was quantified on a Molecular Dynamics PhosphorImager with ImageQuant software.

Linker substitution mutagenesis. Substitution mutants were introduced within CC10 upstream sequences by the method of Kunkel et al. (24). A 2.0-kb *Acc*I-*Bam*HI fragment was liberated from plasmid pCC10CAT-2300 (39) and subcloned into M13 mp19. Mutagenic oligonucleotides were 40 bases in length. Each oligonucleotide encoded an *Acc*I restriction site in the context of a 10-base substitution (CC10-1 through CC10-9) which was flanked on each side by 15 unaltered bases (Fig. 2). The identities of mutants were confirmed by dideoxy DNA sequence analysis, and CC10-chloramphenicol acetyltransferase (CAT) expression cassettes were subcloned into the multiple cloning site of pGEM7Z (Promega). A substitution/deletion mutant lacking region I was generated by replacement of the 235-bp *Acc*I fragment from CC10-8 with the 195-bp *Acc*I fragment from CC10-5 (Fig. 2).

Expression assays. Transient CAT expression assays were performed with the human lung adenocarcinoma cell line

A.



B.

Construct Designation	Substituted Sequence	Relative CAT Activity / %
wild-type		100
CC10-1	(-220) ATGTCGACAT (-210)	53
CC10-2	(-162) GTCGACATCA (-153)	98
CC10-3	(-149) ACTAGTCGAC (-140)	104
CC10-4	(-136) ACGTCGACTA (-127)	104
CC10-5	(-123) GCGTCGACGA (-114)	44
CC10-6	(-110) ACTGTCGACG (-101)	48
CC10-7	(-97) GTCGACTGCT (-88)	77
CC10-8	(-84) TGGTCGACAT (-75)	60
CC10-9	(-42) TACGTCGACT (-33)	93
CC10-5/8	(-123) GCGTC(-118/-80)GACAT (-75)	27

FIG. 2. Site-specific mutagenesis of the upstream CC10 promoter region. (A) Locations of substitution mutations introduced into the promoter proximal region of the CC10 gene; (B) substituted sequences and relative CAT expression for each of the mutations. Transfections were performed three to nine independent times, using different preparations of plasmid DNA. Standard deviations were typically in the range of 5 to 20% of activity. Boundaries of sequences substitutions are indicated in parentheses.

H441 as described previously (39). Mixtures of plasmid DNA containing 5 pmol of test plasmid and 2 pmol of the internal control plasmid pCMVβ were cotransfected into H441 cells. CAT and β-galactosidase assays were performed on cell extracts 72 h posttransfection. The percentage of CAT acetylation was measured by ascending thin-layer chromatography followed by quantitation of the acetylated and nonacetylated products with a Molecular Dynamics PhosphorImager. CAT activities were normalized to β-galactosidase expression and reported as percentage of CAT expression relative to the value for the wild-type CC10 construction.

EMSA. All oligonucleotides used as probes or competitors for EMSA were synthesized on an Applied Biosystems 394 DNA/RNA synthesizer (Fig. 3). Probes were end labeled with [γ-^{32}P]ATP by polynucleotide kinase. Labeled oligonucleotides were then end filled with Klenow enzyme, since most sequences contained GC overhangs for other purposes. Labeled end-filled probes were then purified on Schleicher & Schuell Elutips. Sources of the oligonucleotide sequences are as follows: CC10 sequences are from the rat sequence (16); the HNF-3 binding site TTR is positions −111 to −90 of the mouse transthyretin promoter (6); the octamer binding site is from Ponce et al. (35), and the oligonucleotide was a gift from Elvira Ponce; the AP-1 binding site is positions +15 to +27 of the surfactant B gene (33); the eG HNF-3 binding site is from the albumin enhancer (29), and the oligonucleotide was a gift from Kenneth Zaret. Region I oligonucleotides with point mutations (see Fig. 10) contained wild-type CC10 sequence from −139 to −70 except for the following substitutions (shown in lowercase): probe B,C was (−121) AgTTGaTTgTTGaAT(−107), probe A,C was (−105)GAGA gtACTAA(−95), probe A was (−105)GAGAgtACTAAtTAA cTA(−88), and null was (−121)AgTTGaTTgTTGaATGGAG AgtACTAAGcAgcTA(−88).

EMSAs were performed as described previously (5), with the following modifications. Binding reactions were carried out in a volume of 15 μl containing 200 μg of double-stranded poly(dI-dC) per ml, 10 μg of sheared salmon sperm DNA per ml, 200 μg of bovine serum albumin per ml, 3 mM $MgCl_2$, 75 mM KCl, 10 mM Tris (pH 8.0), 1 mM dithiothreitol, 100 μM phenylmethylsulfonyl fluoride, and 12% glycerol. A total of 25,000 cpm of end-labeled oligonucleotide (specific activity, 10^9 cpm/μg; 50 to 500 pmol) and 1 or 2 μl of nuclear extract (5 to 15 μg of protein) were used in each reaction, and binding was allowed to occur for 15 min at 30°C before the reactions were run at 120 V on 4% nondenaturing acrylamide gels (80:1 acrylamide/bisacrylamide ratio) in 50 mM Tris–380 mM glycine–2 mM EDTA. The HeLa extract lanes of Fig. 8C were run on an 8% acrylamide gel.

When competitions were performed with an excess of cold oligonucleotide, competitor was preincubated with the nuclear extract for 10 min at 30°C prior to the addition of the labeled probe; the reactions were then allowed to incubate for an additional 15 min at 30°C before electrophoresis. When supershifts were performed with HNF-3 antibodies (1 μl), antiserum to Oct-1 (0.5 μl), or antisera to AP-1 family members (1 μl), nuclear extract was preincubated with antibody for 20 min at room temperature prior to the addition of labeled probe; the reactions were then allowed to incubate for an additional 30 min at 30°C before electrophoresis. The anti-Oct-1 antiserum was a gift from Robert Roeder. Affini-

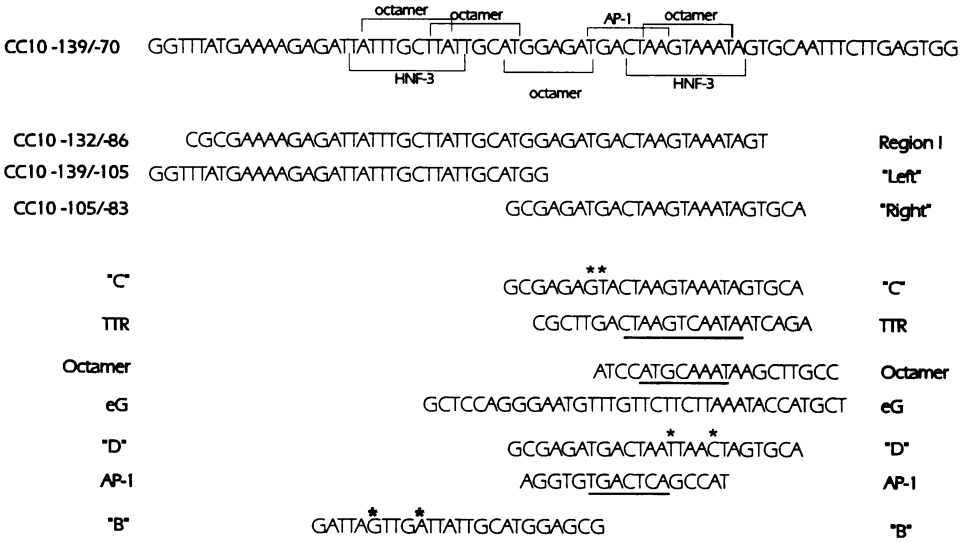


FIG. 3. CC10 region I, HNF-3, octamer, and AP-1 binding site probes and competitors for EMSA. CC10 oligonucleotides spanning −139 to −70 and −132 to −86 are complete region I probes. Factor binding sites predicted from consensus sequences are indicated by brackets on the CC10 −139/−70 sequence. The positions at which sequence changes were made in oligonucleotides B, C, and D relative to wild-type CC10 region I are indicated by asterisks. Locations of consensus factor binding sites in the TTR (HNF-3), octamer, and AP-1 oligonucleotides are indicated by underlines. See Materials and Methods for sequence sources.

ty-purified antibodies to HNF-3α and HNF-3β were a gift from Robert Costa. Antisera to the AP-1 factors were a gift from Rodrigo Bravo (for a description of these antibodies, see reference 23).

RESULTS

Stimulation of in vitro transcription in H441 extracts by CC10 upstream sequences. Two distinct DNA binding sites were identified on the CC10 gene fragment spanning −314 to +59 by DNase I protection analysis using H441 and HeLa nuclear extracts. Region I was centered at −110, and region II was centered at −220. The locations of these sites correlated with transfection data from H441 cells, which suggested that Clara cell-specific regulatory sequences reside on fragments spanning −314 to −165 and −165 to +59 of the CC10 gene (39). DNA fragments with various amounts of CC10 upstream promoter sequence (Fig. 1A) were transcribed in a cell-free system utilizing H441 or HeLa cell nuclear extract. H441 extracts supported transcription from all CC10 promoters tested, whether they contained 833 (not shown), 314, 165, or 97 bp of upstream promoter sequence (Fig. 1B). However, H441 extract produced twice as much runoff transcript from the CC10 −314 and −165 promoters, which contain region I, as from the CC10 −97 promoter, which lacks region I (Fig. 1B).

Although HeLa cells did not support transient expression of the CC10 promoter in vivo (39), HeLa extract did support transcription of CC10 promoters in vitro. However, unlike H441 extract, HeLa extract transcribed all CC10 promoters to an equivalent extent. The ratio of transcriptional activity of H441 extract relative to HeLa extract was about 2 for the region I-containing promoters but only 0.7 for the −97 promoter which lacks an intact region I (Fig. 1B).

As a control for variations in the transcriptional competency of the nuclear extracts, the H441 and HeLa extracts were also assayed on a mouse β-major globin promoter bearing no mammalian sequences upstream of TATA. HeLa extract transcribed equal amounts of RNA from the minimal globin promoter, CC10 −314 promoter, or CC10 −97 promoter, confirming that CC10 upstream sequences do not stimulate transcription in HeLa extract (Fig. 1C; compare HeLa lanes). In contrast, H441 extract transcribed 2.8 or 2.4 times as much transcript from the CC10 −314 promoter compared with the CC10 −97 or globin minimal promoter (Fig. 1C; compare H441 lanes). In several replicates of this experiment, the ratio of H441 to HeLa transcription was typically greater than 2.0 from promoters that contain CC10 region I but only 0.4 to 0.7 from promoters lacking region I.

The transcriptional activation conferred by upstream regions of the CC10 promoter in reactions with H441 extract provided evidence for the existence of a positive regulatory element between −165 and −97 of the CC10 promoter, consistent with the location of region I. At least one essential member of the set of nuclear proteins required to confer stimulation from this element was apparently absent in HeLa extracts. Although the region I-mediated in vitro stimulatory activity of some H441 extract preparations on the −165 CC10 promoter was as great as fivefold (not shown), other preparations failed to show stimulatory activity. We believe that at least one component of the activity required to stimulate in vitro transcription of the CC10 promoter is very labile, since H441 extracts rapidly lost stimulatory, but not basal, transcriptional activity upon storage at −135°C (not shown). Since the region I-mediated stimulation observed in the in vitro transcription assay was not great enough to undertake detailed deletion analysis, the importance of region I in CC10 expression was corroborated by extending the previous transfection experiments of H441 cells with CC10-driven CAT expression constructs (39).

Transient expression assays in H441 cells demonstrate the importance of region I for CC10 expression. To determine the regulatory properties of protein-binding motifs within the CC10 promoter, a series of 10 base substitutions was made and the influence of these changes on promoter function was monitored in transient expression assays (Fig. 2). Five substitution mutations resulted in decreased promoter strength, as monitored by CAT expression. Mutations

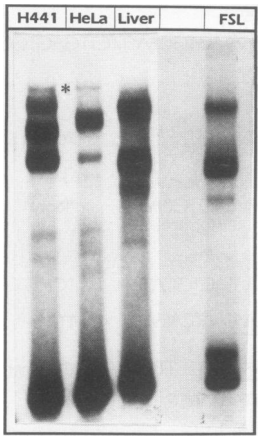


FIG. 4. Evidence that the EMSA pattern of nuclear proteins binding to CC10 region I is unique in H441 extracts. EMSAs were performed with the −132 to −86 region I probe and H441, HeLa, or rat liver extract or the −139 to −70 probe and FSL extract as described in Materials and Methods. The slow-migrating band in the H441 and HeLa patterns is indicated by an asterisk. The FSL pattern was produced on a different gel than that used for the other lanes; the FSL bands were aligned with the liver bands, since the major liver and FSL bands always comigrated in other EMSA experiments (not shown).

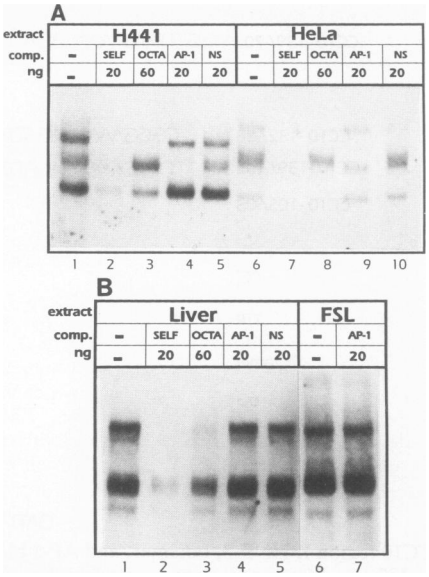


FIG. 5. Challenge by consensus AP-1 and octamer oligonucleotides in region I EMSA, revealing differential competition in each nuclear extract. EMSAs were performed as described in Materials and Methods with the −132 to −86 (H441, HeLa, and liver) or −139 to −70 (FSL) probes after preincubation with the indicated amounts of nonlabeled competitor DNA. NS, nonspecific competitor (*Hae*III-digested pUC18) which could be included in the reactions at up to 80 ng with no competition of major bands (not shown). Sixty nanograms of octamer competitor was used because smaller amounts do not compete for major H441 or liver complexes (see Results).

CC10-5, CC10-6, CC10-7, and CC10-8 involved sequence substitutions within region I and resulted in reduction of CAT expression to 44, 48, 77, and 60%, respectively, of the wild-type level. Additionally, mutation CC10-1, which eliminated the majority of sequences comprising region II, resulted in a 47% reduction in promoter strength (Fig. 2).

To determine the influence of the entire region I element on promoter strength, reporter gene expression was monitored following transfection of the H441 cell line with a substitution/deletion mutation of the −314 to +57 fragment which lacks region I (CC10-5/8; Fig. 2). Expression of CAT from the CC10-5/8 mutation was reduced to 27% of wild-type levels. However, reporter gene expression from the CC10-5/8 mutation was still H441 specific, as no CAT expression was detected after transfection into HeLa cells (not shown).

EMSAs of region I. To investigate the interactions of potential regulatory proteins with region I of the CC10 promoter, EMSAs were performed with H441, HeLa, rat liver, and FSL nuclear extracts. Liver extract was included as a source of factors such as HNF-1 and HNF-3, which have also been reported to be present in the lung (28, 31). The FSL extract (82 days of gestation) provided a source of nuclear proteins from alveolar epithelial and pulmonary endothelial cells; this tissue does not contain a high percentage of Clara cells. The DNA probes that were used in these assays are diagrammed in Fig. 3 and 10. EMSA patterns generated with each of these extracts by using CC10 region I as a probe are shown in Fig. 4. The H441 pattern had four bands, including three strong complexes and a fourth weaker complex (marked with an asterisk in Fig. 4), which was occasionally visible only upon overexposure of the autoradiograms. HeLa extract produced a three-band pattern which was different from that of H441. Although equivalent amounts of nuclear protein were used in all assays, the HeLa bands were considerably less intense than those formed in H441, liver, or FSL extracts. The fastest- and slowest-mobility bands in the H441 and HeLa EMSA patterns comigrated. Rat liver and FSL extracts yielded an identical pattern with two major bands, neither of which comigrated with H441 or HeLa complexes. FSL extracts from early-, middle-, and late-gestation fetuses gave similar results (not shown).

To further characterize protein-DNA interactions on CC10 region I, a series of probes was designed to isolate potential protein binding sites within this DNA (Fig. 3). The region was divided into two parts. The left (upstream) side, from −139 to −105, contains a pair of overlapping octamer (7/8 and 6/8 match) sites; within these sites is an 8/11 match to the consensus HNF-3 binding site (9). The right (downstream) side, from −105 to −83, contains an AP-1 site which overlaps with an 10/11 match to the HNF-3 consensus site. The latter sequence includes within it a 6/8 match to the octamer binding site. Point mutations were introduced into the left side to yield probe B, which still contains one octamer site but has an impaired upstream HNF-3/octamer site. Mutations were introduced into the right side to yield probe C, which has an intact HNF-3/octamer motif but an impaired AP-1 site. Finally, probe D contains an intact AP-1 site but an impaired HNF-3/octamer site. These CC10 sequences, along with others containing consensus binding sites for AP-1, octamer, and HNF-3 transcription factors, were used to investigate the binding of factors in each nuclear extract to CC10 region I.

AP-1 binding to CC10 region I. Although all region I EMSA patterns were stable to competition with an excess of nonspecific competitor such as pUC18 DNA fragments (Fig. 5, NS lanes), a consensus AP-1 site was an efficient competitor of the central bands in region I EMSAs in assays using H441 and HeLa extracts (Fig. 5A, lanes 4 and 9). In contrast, the region I complexes formed in EMSAs using liver or FSL

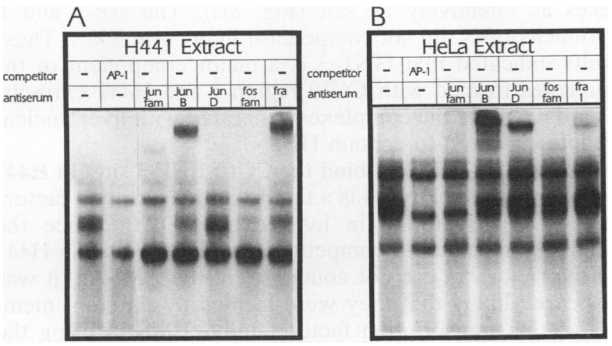


FIG. 6. Evidence that antisera to AP-1 family members supershift H441 and HeLa extract/region I complexes which were shown to contain AP-1 via oligonucleotide competition. EMSAs were performed as described in Materials and Methods with the −132 to −86 region I probe and the indicated extracts. Each extract was preincubated with AP-1 oligonucleotide or antiserum for 20 min prior to the addition of the −132 to −86 probe, as indicated.

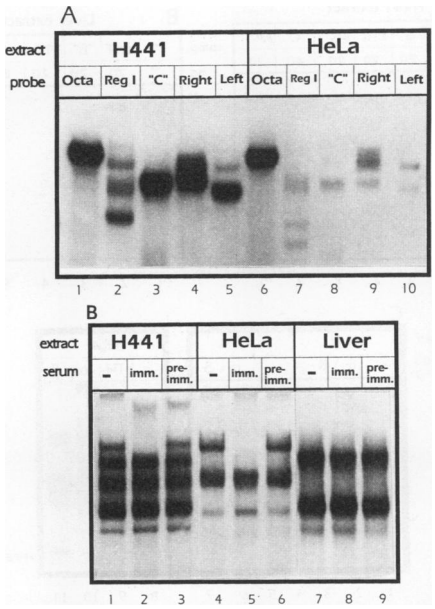


FIG. 7. Evidence that Oct-1 binds to CC10 region I weakly in HeLa and H441 extracts but does not bind at all in liver extracts. (A) EMSAs were performed as described in Materials and Methods with the indicated probes and extracts; the region I probe spanned −132 to −86. (B) EMSAs were performed with the −132 to −86 region I probe and the indicated extracts. Some reactions were incubated with anti-Oct-1 antiserum or preimmune serum for 10 min prior to the addition of the −132 to −86 probe, as indicated. Twice as much HeLa extract was used in this experiment to match the signal in the HeLa EMSAs to that of H441 and liver extracts.

extracts were not sensitive to competition by the AP-1 oligonucleotide (Fig. 5B, lanes 4 and 7). The central AP-1 complexes in H441 and HeLa extract were sometimes seen as a poorly defined doublet, the lower part being stable to competition with the AP-1 oligonucleotide (see Fig. 8A, lane 7, for H441 and Fig. 6B, AP-1 lane, for HeLa). However, since the presence of this lower band of the doublet was erratic, we have been unable to identify it.

A preliminary Northern (RNA) blot analysis indicated that H441 cells contain mRNAs for *junB* and *fra1* but not for c-*jun*, *junD*, or c-*fos* (not shown). To identify which members of the AP-1 transcription factor family were binding to region I in H441 extract, antisera specific for JunB, JunD, Fra1, all Jun family members, or all Fos family members were used to supershift the AP-1-containing EMSA complexes. Figure 6A shows that in H441 extract, antisera to JunB and Fra1 produced supershifted complexes. The JunD antiserum failed to react with the H441 complexes. The Jun family and Fos family antisera prevented formation of the AP-1 complex with H441 extract and the region I probe but did not produce supershifts.

It has been reported that the AP-1 activity in HeLa extracts is predominantly due to c-Jun homodimers (8). However, when we used the AP-1 family antisera to classify the AP-1 activity in HeLa extract/region I EMSAs (Fig. 6B), we found that the JunB and JunD antisera supershifted the HeLa AP-1 complex. As in the H441 reaction, the Jun family antiserum did not form a supershift but did prevent formation of the AP-1 complex. These data indicate that JunB and JunD, and probably c-Jun, bind region I in HeLa extract to some extent. In contrast to the H441 result, the Fos family and Fra1 antisera reacted very weakly with the HeLa/region I complex.

Oct-1 binding to CC10 region I. When a consensus octamer site was used as competitor in region I EMSAs, some complexes generated with H441, HeLa, or liver extracts were sensitive (Fig. 5). The octamer site oligonucleotide competed for all of the HeLa complexes which were not sensitive to the AP-1 oligonucleotide (Fig. 5A, lane 8). However, a competitor titration revealed that the consensus octamer oligonucleotide was only one-seventh as effective as the region I oligonucleotides in reducing the level of the major H441 and liver EMSA complexes (not shown). The upper major bands in H441 and liver were most sensitive to octamer competition (Fig. 5A, lane 3; Fig. 5B, lane 3); note that three times the mass of octamer oligonucleotide was used in these experiments relative to the other competitors.

To investigate the possibility that H441 cells might contain a lung-specific octamer factor, a consensus octamer site was used as probe in EMSAs with H441 or HeLa extracts (Fig. 7A). H441 extract yielded a single complex with this probe. A complex of identical mobility was produced by HeLa extract with the octamer probe (compare lanes 1 and 6). Since HeLa cells are known to contain the ubiquitous transcription factor Oct-1 as their only octamer-binding activity (34), this comigration indicated that H441 cells also contain Oct-1 as their only identifiable octamer binding activity. None of the major complexes produced with H441 proteins and various partial CC10 probes (lanes 3 to 5) migrated with the bona fide Oct-1 complex from HeLa extract (lane 6). This finding suggested that in H441 extract, Oct-1 does not bind strongly to region I. However, the weaker complex produced with H441 extract and the region I probe (asterisk in Fig. 4) did appear to migrate with the complex obtained with the octamer consensus probe, suggesting that this complex could be due to Oct-1 binding. The significance of the observation was uncertain, since the lengths of the region I and octamer oligonucleotides were different in this experiment.

To definitively establish the role of the Oct-1 protein in the complexes formed with region I probes, antiserum raised against Oct-1 was used in EMSAs with region I probes and H441, HeLa, or liver extract (Fig. 7B). Preincubation of the nuclear extracts with anti-Oct-1 antiserum prior to probe addition prevented formation of the slowest-mobility com-

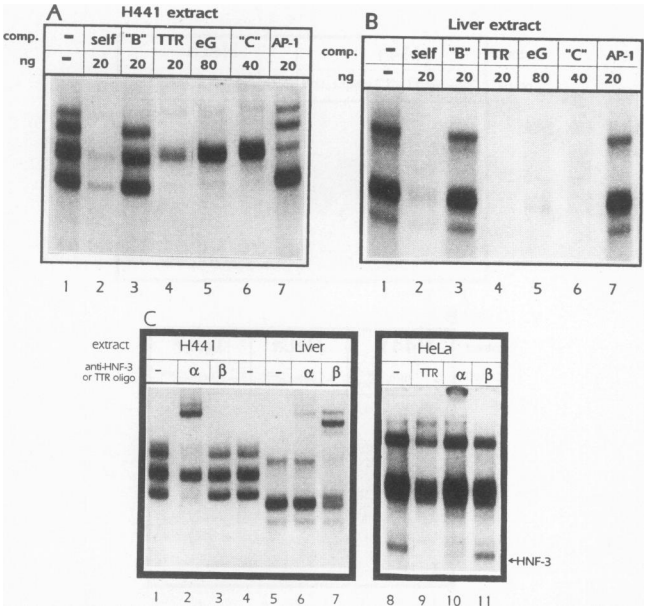


FIG. 8. Binding of HNF-3 family members to CC10 region I. EMSAs were performed as described in Materials and Methods with the −132 to −86 probe and H441 (A) or rat liver (B) nuclear extract after preincubation with the indicated amounts of nonlabeled competitor DNA. (C) Reactions containing the indicated extracts were preincubated with the TTR oligonucleotide or antibodies to HNF-3α or HNF-3β for 20 min prior to addition of the −132 to −86 probe, as indicated. The difference in migration of the HNF-3α complex and supershift in the HeLa lane relative to the H441 lane was due to the difference in porosity of the gels (see Materials and Methods); the HNF-3α band in HeLa extract is indicated. Twice as much HeLa extract was used in this experiment to match the signal in the HeLa EMSAs to that of H441 and liver extracts.

plex in H441 and HeLa extract EMSAs (lanes 2 and 5). This effect was not seen with preimmune serum (lanes 3 and 6); therefore, this complex must contain Oct-1. Interestingly, none of the partial region I probes were bound by Oct-1 in H441 or HeLa extract (Fig. 7A and data not shown), indicating that Oct-1 needed an intact region I to bind under these conditions. Neither immune nor preimmune serum prevented the formation of any complexes observed in liver extract (lanes 8 and 9).

HNF-3 binding to CC10 region I. When HNF-3 binding sites were used as competitors in region I EMSAs, complexes in H441 and liver extracts were found to be sensitive (Fig. 8). The HNF-3 binding sites used as competitors were the strong HNF-3 binding site present in the mouse transthyretin promoter (TTR [6]) and an HNF-3 binding site from the albumin enhancer (eG [29]), which binds HNF-3 about one-fourth as well as TTR does. In reactions with H441 extract (Fig. 8A), TTR, eG, and the CC10 downstream HNF-3 site (C) competed equally for the major upper and lower complexes. As noted above (Fig. 5), the AP-1 site competed for the central band only. Competition with the CC10 B oligonucleotide was included in this experiment as a nonspecific control since this sequence contains no AP-1 or HNF-3 binding sites; as expected, B failed to affect the major H441 complexes. However, note that B contains a 6/8 and TTR contains a 5/8 match to the octamer sequence. It is therefore not surprising that both of these DNAs partially competed for the weaker Oct-1 complex. In assays with liver extract, the TTR, eG, and C DNAs competed for all complexes as effectively as self (Fig. 8B). The AP-1 and B oligonucleotides did not compete for liver complexes. These results indicated that HNF-3 is a major component of the complexes formed with CC10 region I and H441 extracts. Essentially all of the complexes generated with liver nuclear extracts appeared to contain HNF-3.

Distinct HNF-3 factors bind the CC10 HNF-3 sites in H441 and liver extracts. HNF-3 is a family of transcription factors which was first found in liver extracts (28). Since the complexes sensitive to competition by HNF-3 sites in H441 and liver extracts did not comigrate in the EMSAs, it was considered likely that they were formed by different members of this transcription factor family. EMSAs using the TTR and C probes with H441 and liver extract were performed; comparison with published HNF-3 EMSA patterns (28) suggested that HNF-3α was binding region I in H441 extract and HNF-3β was binding region I in liver extract (not shown). To verify this finding, antibodies specific to HNF-3α and HNF-3β were used to supershift H441 or liver complexes formed with the region I probe. The results in Fig. 8C confirm that the upper and lower major complexes generated with H441 extract and the region I probe contained HNF-3α, as they were both supershifted by the antibody to HNF-3α (lane 2). These same complexes were unreactive with the HNF-3β antibody (lane 3). In contrast, region I complexes generated in liver extract were predominantly reactive to the HNF-3β antibody (lane 7) and showed greatly reduced reactivity with the HNF-3α antibody (lane 6).

The amount of H441 or liver extract used in EMSAs with the region I probe dictated the proportion of each HNF-3 complex produced in the reaction. At very low extract concentrations, only the faster-mobility HNF-3 complex formed. However, at high extract concentrations, the slower-mobility band predominated. Indeed, when very high concentrations of liver extract were used, only the slower-mobility complex formed. These results (not shown) suggest that in the faster-mobility HNF-3 complex, only one of the two region I HNF-3 binding sites is occupied, whereas in the slower-mobility complex, both HNF-3 binding sites are occupied. This assertion is supported by the fact that the degenerate HNF-3 binding site contained within the octamer probe competed to a much greater extent for the slower-mobility HNF-3 complex than for the faster-mobility HNF-3 complex (Fig. 5A, lane 3; Fig. 5B, lane 3).

HeLa extract has not been reported to be a source of HNF-3. However, the identity of the fastest-migrating complex in HeLa/region I EMSAs had not yet been established. Since this band comigrated with the faster-mobility HNF-3α complex of H441 extract (Fig. 4), it was possible that this band contained HNF-3α. Competition of a HeLa/region I EMSA with the consensus HNF-3 site (TTR) efficiently competed for this complex (Fig. 8C, lane 9). Additionally, inclusion of the antibody to HNF-3α, but not the antibody to HNF-3β, supershifted this complex, indicating that this band is due to HNF-3α binding (lane 10). Notice that for lanes 8 to 11 of Fig. 8C, a higher percentage polyacrylamide gel was used (8%) than in the experiments shown in Fig. 4 and Fig. 8, lanes 1 to 7 (4%).

Can AP-1 and HNF-3 bind region I simultaneously? The AP-1 and downstream HNF-3 binding sites present in region I overlap, and it seemed unlikely that both AP-1 and HNF-3 would bind these sites simultaneously. To test this inference, the region I probe which contains only these two binding sites (right probe) was used in EMSAs with either the CC10 HNF-3 site or the CC10 AP-1 site (contained on oligonucleotides C and D, respectively; see Fig. 3) as competitor. The

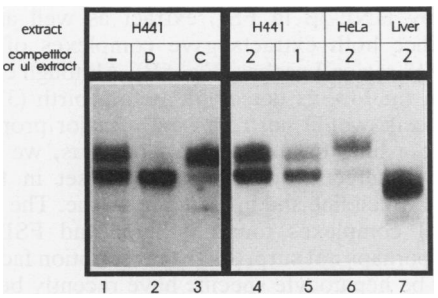


FIG. 9. Ability of HNF-3 or AP-1 to bind to the CC10 right probe. EMSAs were performed as described in Materials and Methods with the CC10 right probe and the indicated extracts. Lanes 1 to 3 contained 1 μl of H441 extract; lanes 2 and 3 were preincubated with 20 ng of the indicated nonlabeled competitor DNAs. Lanes 4 to 7 contained 1 or 2 μl of extract as indicated and no competitor DNA.

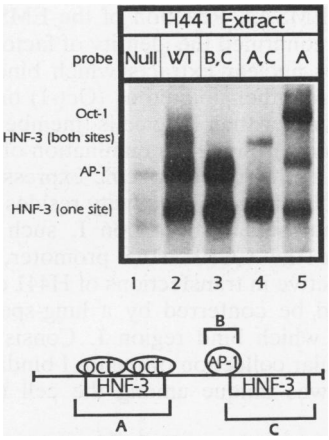


FIG. 10. H441 extract EMSA patterns on region I probes containing point mutations in factor binding sites. EMSAs were performed as described in Materials and Methods with H441 extract and either wild-type region I (WT) or the point-mutated versions of region I diagrammed in the simplified cartoon at the bottom. Probe names reflect the presence of the factor binding sites indicated in the cartoon. Sequences for the probes are given in Materials and Methods. The proteins whose binding generates the four complexes observed with the wild-type region I probe are indicated at the left.

H441 extract EMSA pattern with the right probe consisted of two bands (Fig. 9, lane 1). The slower complex was due solely to AP-1 binding, as it was sensitive to competition with either the CC10 AP-1 site (lane 2) or the consensus AP-1 site (not shown) but not the CC10 HNF-3 site (lane 3). The faster complex was due solely to HNF-3α, as it was sensitive to competition with the CC10 HNF-3 site (lane 3) or TTR (not shown) but not the CC10 AP-1 site (lane 2). The faster complex was also supershifted by anti-HNF-3α antibody (not shown).

As expected, two complexes were observed with the right probe in HeLa extract (Fig. 9, lane 6), but only a single complex was observed in liver extract (lane 7). Competition with appropriate oligonucleotides has shown the slower band in EMSAs with HeLa extract and the right probe to be the AP-1 complex and the faster band to be the HNF-3 complex (not shown). The region I AP-1-containing complexes from H441 and HeLa extracts did not comigrate, supporting the conclusions from the antibody analysis that these complexes contain distinct AP-1 activities. In contrast, the H441 and HeLa HNF-3 complexes did comigrate, indicating that they contain the same factor, namely, HNF-3α. The intensities of the faster complexes in lanes 4 and 6 of Fig. 9 indicated the relative amount of HNF-3α in H441 and HeLa extracts, since the extracts used had roughly equivalent protein concentrations (7.4 and 6.5 mg/ml for H441 and HeLa, respectively). The EMSA with liver extract and the right probe showed a strong but diffuse band. This band was not sensitive to competition by the CC10 AP-1 site (not shown). Competition and antibody supershift experiments proved that this band was composed of predominantly HNF-3β, although some HNF-3α was also able to bind (not shown); HNF-3α binding accounted for the indistinct upper border of this band.

Do proteins from H441 nuclear extract bind region I probes which contain point mutations in factor binding sites? Once the factors which bound region I in the Clara cell-like H441 extract were identified, mutations were introduced into this region to demonstrate that disruption of one or more binding sites resulted in the predicted loss of factor binding. The mutations were designed to contain the minimal number of base substitutions necessary to recreate the loss of binding sites caused by the linker scanning mutations CC10-5, -6, and -7. For the purpose of describing these probes, the upstream HNF-3/octamer site was called A, the central octamer/AP-1 site was called B, and the downstream octamer/HNF-3 site was called C (see simplified cartoon in Fig. 10). Probe B,C contained only the AP-1 and downstream HNF-3/octamer binding sites (similar to CC10-5). Incubation of this probe in H441 extract generated two complexes (Fig. 10, lane 3); one comigrated with the AP-1 complex and the other migrated with the fast (monomeric) HNF-3 complex of the wild-type −139/−70 probe (lane 2). Probe A,C contained the upstream and downstream HNF-3/octamer binding sites (similar to CC10-6). This probe bound two complexes (lane 4), which comigrated with the fast (monomeric) and slow (dimeric) HNF-3 complexes of the wild-type probe. Probe A contained only the upstream HNF-3 and octamer binding sites (similar to CC10-7). Two of the complexes generated with this probe (lane 5) comigrated with the fast HNF-3 and Oct-1 complexes of the wild-type probe. A weaker third band comigrated near the AP-1 complex of the wild-type probe but was not sensitive to AP-1 competition (not shown). We have not yet determined the identity of this band; it does not correspond to any of the complexes described above. The null probe contained point mutations which would disrupt binding sites A, B, and C (similar to CC10-5/8). Although this probe is 70 bp in length, just nine nucleotide substitutions abolished all specific protein binding to this mutated version of CC10 region I (lane 1).

DISCUSSION

Our earlier analysis of the protein-DNA interactions at the CC10 promoter demonstrated that region I, located 79 to 128 bp upstream of the transcriptional start site, bound nuclear proteins in a cell-type-dependent manner and may therefore be involved in the Clara cell-specific expression of CC10 (39). In the present study, in vitro transcription results and transient expression experiments utilizing linker scanning mutations of region I are consistent with a role for this element in regulation of the CC10 gene in H441 cells. The identities of the transcription factor families which bind to region I were determined by using CC10 sequences and consensus transcription factor binding sites as probes and

competitors in EMSAs. Reaction of the EMSA complexes with antibodies confirmed the identity of factors from H441, HeLa, and liver nuclear extracts which bind to region I. These factors are either ubiquitous (Oct-1) or expressed in many cell types other than lung cells (members of the AP-1 and HNF-3 families). Since the elimination of region I does not completely abolish H441-specific expression in our assays, it is possible that lung specificity resides, entirely or in part, in elements outside of region I, such as region II. However, since the CC10 −165 promoter, which lacks region II, was active in transfections of H441 cells (39), lung specificity could be conferred by a lung-specific array of generic factors which bind region I. Consistent with this idea, the particular collection of region I binding proteins in H441 extracts was unique among the cell types that we tested.

Which factors interact with CC10 region I? Oct-1 was found to bind weakly to CC10 region I in H441 and HeLa extracts, giving rise to the slowest-migrating complex formed on the region I probe with either of these extracts (Fig. 7B, lanes 1 and 4). This weak Oct-1 binding (Fig. 7A; compare Oct-1 binding to the consensus octamer probe in lane 1 versus the region I probe in lane 2) can be explained by the presence of four degenerate octamer sites within region I. It has been demonstrated that Oct-1 can bind degenerate octamer sequences, although an imperfect match to consensus reduces affinity (1). Oct-1 did not bind any of the partial region I probes (Fig. 7A), so initially we could not determine which octamer binding site was occupied by Oct-1 in H441 nuclear extracts. However, in the experiment in Fig. 10, Oct-1 was found to bind strongly to a point-mutated probe lacking the downstream octamer sites (probe A, lane 5), suggesting that the two upstream sites within region I are most important for Oct-1 binding. Oct-1 bound probe A much more strongly than did the wild-type probe (Fig. 10; compare lanes 2 and 5). Interestingly, the point-mutated region I probe A,C does not bind Oct-1 (Fig. 10, lane 4), despite the fact it contains the same upstream octamer binding sites as does probe A. This situation mirrors that of the left probe (Fig. 7A, lane 5), which also contains both upstream octamer binding sites but does not bind Oct-1. Further studies will be required to determine how protein-protein interactions within region I influence the binding of Oct-1 to this DNA sequence.

The initially confusing observation that competition with the consensus octamer site eliminated some HNF-3 complexes in H441, HeLa, and liver EMSAs at high oligonucleotide concentrations (Fig. 5A, lanes 3 and 8; Fig. 5B, lane 3) is explained by the fact that our octamer probe contained a 7/11 match to the HNF-3 consensus site. The EMSA reactions employing the anti-Oct-1-specific antisera demonstrated that the complexes which were only moderately sensitive to competition with the octamer oligonucleotide did not contain Oct-1.

Complexes formed with H441, liver, or HeLa extract on region I were sensitive to competition with HNF-3 binding sites in an EMSA (Fig. 8). Antibodies specific to HNF-3β supershifted region I complexes made with liver extract (Fig. 8C, lane 7) but were unreactive in H441 or HeLa assays (lanes 3 and 11). Antibodies specific to HNF-3α supershifted region I complexes with all three extracts (lanes 2, 6, and 10), although the liver extract/HNF-3α supershift was much weaker than the HNF-3β supershift.

The relatively simple EMSA pattern formed with the region I probe and rat liver extract was the result of HNF-3β binding. We believe that the region I HNF-3 sites are occupied by HNF-3β in FSL extract as well as in liver extract, since both extracts gave complexes of identical mobility with region I probes (Fig. 5B). Although expression of CC10 in the lung is detectable before birth (37), CC10-producing cells would not represent a major proportion of the cells in a lung tissue preparation. Thus, we were not expecting to detect an identical factor set in the Clara cell-like H441 cell line and in fetal lung tissue. The similarity of region I complexes found in liver and FSL nuclear extracts is perhaps not surprising. Transcription factors once thought to be hepatocyte specific have recently been found to be present in a variety of tissues derived from primitive gut endoderm, such as the lung and the intestine (27). It will be of interest to compare transcriptional regulation in these two organs of foregut origin, as it has been suggested that their biochemical maturation processes may be analogous (10).

Only HNF-3α binds region I in H441 and HeLa extracts, in contrast to liver extracts, in which HNF-3β binding strongly predominates. EMSAs performed with partial region I probes which contain only the upstream or downstream HNF-3 site demonstrated that a single major complex was observed in H441 extract (Fig. 7A, lanes 3 and 5). Antibody supershifts have confirmed that these complexes contain HNF-3α (not shown). Indeed, we have been able to detect little, if any, HNF-3β in H441 extracts with use of a variety of probes (not shown). As in the case of liver and FSL extracts, titrations of H441 extract in region I EMSAs suggested that both HNF-3 sites are occupied at high extract concentrations; only the fastest-migrating HNF-3α complex formed at low extract concentrations, while the slower-migrating HNF-3α complex predominated at high extract concentrations (not shown).

The low amount of HNF-3α binding to region I seen with HeLa extract is consistent with the weak HNF-3α signal this extract produces with the TTR probe (not shown). Incubation of high levels of HeLa extract with a region I probe did not generate two bands sensitive to competition with HNF-3 binding sites. EMSAs employing HeLa extract and the partial region I probes showed low but reproducible levels of HNF-3α binding to the right probe (faster band in Fig. 7A, lane 9) but almost no binding to the upstream HNF-3 site on the left probe (lane 10). This finding indicates that the upstream CC10 HNF-3 site probably has a weaker affinity for HNF-3α than does the downstream site and would thus be the second site to be occupied at high extract concentrations in H441 extract. This conclusion is consistent with the fact that the upstream sequence matches the HNF-3 consensus at only 8 of 11 positions, while the downstream site present on the C probe matches at 10 of 11 positions.

AP-1 was found to bind to CC10 region I in H441 and HeLa cell nuclear extracts. Antisera specific to two AP-1 family members, JunB and Fra1, were able to supershift the H441 extract/region I AP-1 complex. These data agree with our observation that H441 cells contain mRNAs for *junB* and *fra1* but do not contain messages for c-*jun*, *junD*, or c-*fos* (not shown). Therefore, the AP-1 type that binds region I in H441 extract is most likely JunB/Fra1 heterodimers; however, we have not yet determined whether Fra2 or FosB is present in H441 extract. In HeLa extract/region I EMSAs, antisera specific to JunB, JunD, and all Jun family members reacted strongly with the HeLa AP-1 complex, whereas antisera specific to Fra1 and Fos family members reacted weakly. Therefore, we cannot determine from our experiments the most likely composition of the AP-1 binding to region I in HeLa cells. The differential reaction to anti-AP-1

antisera of AP-1-containing complexes formed on region I with H441 or HeLa extract suggests that these complexes contain different AP-1 proteins. This observation is supported by the fact that the mobilities of these complexes are different in region I EMSAs (Fig. 9; compare slower complexes in lanes 5 and 6).

Use of CC10 right as the labeled probe in EMSAs with H441 and HeLa extracts confirmed that both AP-1 and HNF-3 interact with their binding sites on this DNA. However, since neither band was sensitive to competition by both binding sites, it is clear that AP-1 and HNF-3 cannot bind simultaneously (Fig. 9). Even in experiments using very high concentrations of extract, we cannot force both factors onto this sequence (not shown). This is not surprising, since these protein binding sites overlap (Fig. 3). Interestingly, although liver extract has been shown to contain AP-1 activity (46), liver extract does not appear to contain an AP-1 factor which will bind to the CC10 region I AP-1 site. Since the region I HNF-3 sites are occupied predominantly by HNF-3β in liver extract, as opposed to HNF-3α in H441 and HeLa extracts, we thought that the strong binding of HNF-3β to region I might exclude potential AP-1 binding. However, competition of a liver extract/region I EMSA reaction with an HNF-3 binding site which lacks an AP-1 site did not reveal an AP-1 binding activity (Fig. 8B, lanes 5 and 6). Additionally, we have recently discovered that in nuclear extracts from the A549 cell line, a human lung adenocarcinoma lacking any Clara cell characteristics, HNF-3β or AP-1 can bind to region I (not shown). Therefore, our failure to detect an AP-1 binding activity in liver extract was not simply the result of the presence of HNF-3β.

What is the significance of the factor interactions within region I for CC10 expression? The strongest evidence for the importance of region I in CC10 expression is the drop in activity of linker scan mutation CC10-5/8 in transient assays relative to the activity of the intact parental construct. In CC10-5/8, entirely unrelated sequence has replaced region I, resulting in a fourfold reduction of promoter activity. Smaller sequence substitutions within region I, namely, CC10-5, -6, -7, and -8, reduced activity of the promoter at most twofold. We would expect that the upstream HNF-3 and octamer sites are lost in CC10-5, the AP-1 site is lost in CC10-6, and the AP-1 site and the downstream HNF-3 and octamer sites are missing in CC10-7. To confirm that the loss of these sites caused a loss of factor interactions, we synthesized four new region I oligonucleotides which should also lack the relevant factor binding sites. To simplify interpretation, we recreated the effects of the linker scans by using the smallest number of point mutations expected to eliminate the sites which were disrupted by linker scanning. These new probes (Fig. 10) were called B,C (the analog of CC10-5), A,C (the analog of CC10-6), and A (the analog of CC10-7). The null oligonucleotide, with a total of nine base changes, should not contain any of the binding sites that we characterized and was thus analogous to CC10-5/8. As shown in Fig. 10, each of the point-mutated versions of region I failed to produce at least one of the four factor complexes generated by H441 extracts on wild-type region I; none of these complexes were obtained with the null oligonucleotide. Thus, the loss of transcriptional activity of the CC10-5, -6, -7, and -5/-8 constructions may reasonably be interpreted as loss of the ability to bind some or all of the transcription factors characterized here.

Surprisingly, mutation CC10-8 occurs in a region in which we did not detect protein binding. However, examination of this region reveals a highly AT-rich sequence. Proteins which bind AT-rich DNA sequences are often difficult to detect in the presence of high amounts of some nonspecific competitor DNAs, such as poly(dI-dC), which were used in these experiments (38). For example, HMG I(Y), which binds DNA in the minor groove, has recently been shown to be important in the regulation of the beta interferon promoter; the sequence disrupted in mutation CC10-8 is very similar to the HMG I(Y) binding site in PRD2 of the beta interferon promoter (42). It is therefore possible that other factors bind to the most downstream portion of region I but were not detected in our binding assays.

Specific members of transcription factor families are required for transcriptional activation by region I. Although we have determined that HeLa extract contains three transcription factors capable of binding to region I in vitro, HeLa cells do not support transient expression from the CC10 promoter (39). We did find that the binding of HeLa nuclear proteins to CC10 region I was weaker than binding of H441 or liver proteins (Fig. 5A or 9). In particular, it is possible that HeLa cells do not contain a high enough level of HNF-3α to activate this promoter. Alternatively, since the type of AP-1 from HeLa extracts which binds region I appears to be different from the region I-binding AP-1 activity in H441 extracts, the binding of the incorrect AP-1 to this promoter could have failed to activate or even repressed transcription in vivo (8). It is important to note that we have also examined the region I factor binding set in A549 cells, which do not support CC10 expression in transfection assays. A549 contains a member of each of the transcription factor families (AP-1, Oct-1, and HNF-3) which bind CC10 region I in H441 extracts; however, the HNF-3 type in A549 is HNF-3β (not shown). The inactivity of the CC10 promoter in A549 cells may indicate that the binding of the "incorrect" transcription factors, such as HNF-3β, to region I may serve to negatively regulate CC10 in lung cell types related to the Clara cell, as has been suggested for other genes (see, for example, reference 4).

Importance of region II to CC10 expression. Our initial survey of the upstream sequences of the CC10 promoter via DNase I footprinting and transient transfection indicated that a second region, located around −220, was also important to CC10 transcription; we termed this region II (39). Although the reduction in activity of CC10-1, a linker scanning mutation of region II, confirms that this region is an important component of CC10 expression, stimulatory transcriptional activity was not observed on the −314 promoter relative to the −165 promoter in the in vitro transcription assay (Fig. 1B). Additionally, we have been unable to produce cell-specific EMSA patterns with region II (not shown). The fact that complete removal of region I failed to abolish all transcription from the CC10 promoter may have been due to the positive effect of region II on the CC10-5/8 promoter.

Transcriptional regulation of CC10 and uteroglobin. Both the amino acid sequence and the sequence of the upstream promoter region of CC10 are very similar to those of rabbit uteroglobin (16). In rabbits, uteroglobin is expressed at high levels in the uterus and at low levels in the lung, which is the reverse of the expression pattern seen with CC10 (16). While the functional and evolutionary relationship between CC10 and uteroglobin is not clear, a comparison of the transcriptional regulation of these genes is of interest since both genes are expressed in the lung and uterus. The uteroglobin promoter contains a region located 90 to 130 bp upstream of the start site which is similar to CC10 region I. This region contains two octamer-like motifs which are believed to bind

Oct-1 (41). Additionally, three sequences in this region of the uteroglobin promoter match the HNF-3 consensus at 9 of 11 positions; however, none of these sequences match either of the CC10 HNF-3 sites.

Studies by Misseyanni et al. (32) detected protein binding to this region in uterine (Ishikawa) and nonuterine (HeLa) extracts in a DNase I protection assay. Analysis of subsequent linker scanning mutations suggested that this region of the promoter was important for transient expression in Ishikawa and H441 cells (41). Interestingly, the effects of each mutation were not equivalent in the two cell types. Preliminary EMSAs using the uteroglobin −90 to −140 sequence as a probe failed to demonstrate binding of H441 proteins. Furthermore, this sequence did not serve as an efficient competitor in H441/region I EMSAs, despite the presence of the three degenerate HNF-3 sites (not shown). Recent studies in which the uteroglobin promoter was used to drive reporter expression in transgenic mice demonstrated that animals carrying a transgene with 600 bp of the uteroglobin promoter expressed the transgene in the uterus and lung (7), while animals carrying a transgene with 400 bp of the uteroglobin promoter expressed the transgene in the uterus only (26). These experiments suggest that sequences present between −400 and −600 bp are responsible for the expression of uteroglobin in the mouse lung. Considering these observations, it seems likely that the sequences responsible for lung-specific expression of the rabbit uteroglobin gene in mice are different from the sequences examined in this study.

Region II is highly GC rich and resembles an SP-1 site. Another group has recently shown that novel transcription factors related to SP-1 interact with a site in a similar location on the uteroglobin promoter (15). We therefore challenged an H441 extract region II/EMSA complex with an SP-1 sequence as competitor. However, this sequence did not compete for the H441/region II complex (not shown). It therefore seems unlikely that an SP-1-like factor is interacting with region II of the CC10 promoter in H441 cells. Further study will be required to determine the importance of region II to CC10 transcriptional regulation.

ACKNOWLEDGMENTS

We thank Peter Ciralo and Jacquelyn Huffman for excellent technical assistance, Charlene McWhinney for the FSL nuclear extract, Kenneth Zaret for valuable discussions and the eG oligonucleotide, Elvira Ponce for the octamer oligonucleotide, Robert Roeder for the anti-Oct-1 antiserum, Robert Costa for the HNF-3 antibodies, Rodrigo Bravo for the AP-1 antibodies, Susan White Luse and Clemencia Colmenares for critically reviewing the manuscript, Tom Doetschman and Bruce Moulton for rat tissue, and Harriet Yamamoto for the fetal sheep tissue.

This research was supported by the Program of Excellence in Molecular Biology of the Heart and Lung HL-41496 (P.L.S. and B.R.S.), the American Lung Association RG-173-N (B.R.S.) and the Cystic Fibrosis Foundation (J.A.W.).

REFERENCES

1. **Baumruker, T., R. Sturm, and W. Herr.** 1988. OBP100 binds remarkably degenerate octamer motifs through specific interactions with flanking sequences. Genes Dev. **2**:1400–1413.
2. **Birkenmeier, E. H., B. Gwynn, S. Howard, J. Jerry, J. I. Gordon, W. H. Landschulz, and S. L. McKnight.** 1989. Tissue-specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein. Genes Dev. **3**:1146–1156.
3. **Cao, Z. D., R. M. Umek, and S. L. McKnight.** 1991. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. **5**:1538–1552.
4. **Cereghini, S., M. Raymondjean, A. G. Carranca, P. Herbomel, and M. Yaniv.** 1987. Factors involved in control of tissue-specific expression of albumin gene. Cell **50**:627–638.
5. **Chodosh, L. A.** 1989. DNA protein interactions; mobility shift DNA-binding assay using gel electrophoresis, p. 12.2.1–12.2.10. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, New York.
6. **Costa, R. H., D. R. Grayson, and J. E. Darnell, Jr.** 1989. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α1-antitrypsin genes. Mol. Cell. Biol. **9**:1415–1425.
7. **De Mayo, F. J., S. Damak, T. H. Hansen, and D. W. Bullock.** 1991. Expression and regulation of the rabbit uteroglobin gene in transgenic mice. Mol. Endocrinol. **5**:311–318.
8. **Diamond, M. I., J. N. Miner, S. K. Yoshinaga, and K. R. Yamamoto.** 1990. Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science **249**:1266–1272.
9. **Faisst, S., and S. Meyer.** 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res. **20**:3–26.
10. **Frank, L.** 1989. Preparation for birth, p. 1141–1198. *In* D. Massaro (ed.), Lung cell biology, vol. 41. Marcel Dekker, Inc., New York.
11. **Gail, D. B., and C. J. M. Lenfant.** 1983. Cells of the lung: biology and clinical implications. Annu. Rev. Respir. Dis. **127**:366–387.
12. **Glasser, S. W., T. R. Korfhagen, S. E. Wert, M. D. Bruno, K. M. McWilliams, D. K. Vorbroker, and J. A. Whitsett.** 1991. Genetic element from human surfactant protein SP-C gene confers bronchiolar-alveolar cell specificity in transgenic mice. Am. J. Physiol. **261**:L349–L356.
13. **Gorski, K., M. Carneiro, and U. Schibler.** 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. Cell **47**:767–776.
14. **Hackett, B. P., N. Shimizu, and J. D. Gitlin.** 1992. Clara cell secretory protein gene expression in bronchiolar epithelium. Am. J. Physiol. **262**:L399–L404.
15. **Hagen, G., S. Müller, M. Beato, and G. Suske.** 1992. Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes. Nucleic Acids Res. **20**:5519–5525.
16. **Hagen, G., M. Wolf, S. L. Katyal, G. Singh, M. Beato, and G. Suske.** 1990. Tissue-specific expression, hormonal regulation and 5′-flanking gene region of the rat Clara cell 10 kDa protein: comparison to rabbit uteroglobin. Nucleic Acids Res. **18**:2939–2946.
17. **Howell, B. W., M. Lagacé, and G. C. Shore.** 1989. Activity of the carbamyl phosphate synthetase I promoter in liver nuclear extracts is dependent on a *cis*-acting C/EBP recognition element. Mol. Cell. Biol. **9**:2928–2933.
18. **Izban, M. G., and D. S. Luse.** 1992. The RNA polymerase-II ternary complex cleaves the nascent transcript in a 3′→5′ direction in the presence of elongation factor-SII. Genes Dev. **6**:1342–1356.
19. **Jacob, G. A., S. W. Luse, and D. S. Luse.** 1991. Abortive initiation is increased only for the weakest members of a set of down mutants of the adenovirus-2 major late promoter. J. Biol. Chem. **266**:22537–22544.
20. **Kalina, M., R. J. Mason, and J. M. Shannon.** 1992. Surfactant protein C is expressed in alveolar type II cells but not in Clara cells of rat lung. Am. J. Respir. Cell Mol. Biol. **6**:594–600.
21. **Karpinski, B. A., G. D. Morle, J. Huggenvik, M. D. Uhler, and J. M. Leiden.** 1992. Molecular cloning of human CREB-2—an ATF/CREB transcription factor that can negatively regulate transcription from the cAMP response element. Proc. Natl. Acad. Sci. USA **89**:4820–4824.
22. **Korfhagen, T. R., S. W. Glasser, S. E. Wert, M. D. Bruno, C. C. Daugherty, J. D. McNeish, J. L. Stock, S. S. Potter, and J. A. Whitsett.** 1990. Cis-acting sequences from a human surfactant protein gene confer pulmonary-specific gene expression in transgenic mice. Proc. Natl. Acad. Sci. USA **87**:6122–6126.
23. **Kovary, K., and R. Bravo.** 1991. Expression of different Jun and Fos proteins during the G_0-to-G_1 transition in mouse fibroblasts:

in vitro and in vivo associations. Mol. Cell. Biol. **11**:2451–2459.
24. **Kunkel, T. A., J. D. Roberts, and R. A. Zakow.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
25. **Lacaze-Masmonteil, T., C. Fraslon, J. Bourbon, M. Raymondjean, and A. Kahn.** 1992. Characterization of the rat pulmonary surfactant protein-A promoter. Eur. J. Biochem. **206**:613–623.
26. **Lahoz, E. G., M. S. Lopez de Haro, P. Esponda, and A. Nieto.** 1992. Tissue-specific and hormonally regulated expression of the puromycin N-acetyltransferase-encoding gene under control of the rabbit uteroglobin promoter in transgenic mice. Gene **117**:255–258.
27. **Lai, E., and J. E. Darnell.** 1991. Transcriptional control in hepatocytes—a window on development. Trends Biochem. Sci. **16**:427–430.
28. **Lai, E., V. R. Prezioso, W. Tao, W. S. Chen, and J. E. Darnell, Jr.** 1991. Hepatocyte nuclear factor 3α belongs to a gene family in mammals that is homologous to the *Drosophila* homeotic gene *fork head*. Genes Dev. **5**:416–427.
29. **Liu, J., C. M. DiPersio, and K. S. Zaret.** 1991. Extracellular signals that regulate liver transcription factors during hepatic differentiation in vitro. Mol. Cell. Biol. **11**:773–784.
30. **Lund, J., L. Nordlund, and J. Gustafsson.** 1988. Partial purification of a binding protein for polychlorinated biphenyls from rat lung cytosol: physiochemical and immunochemical characterization. Biochemistry **27**:7895–7901.
31. **Mendel, D. B., L. P. Hansen, M. K. Graves, P. B. Conley, and G. R. Crabtree.** 1991. HNF-1α and HNF-1β (vHNF-1) share dimerization and homeo domains, but not activation domains, and form heterodimers in vitro. Genes Dev. **5**:1042–1056.
32. **Misseyanni, A., J. Klug, G. Suske, and M. Beato.** 1991. Novel upstream elements and the TATA-box region mediate preferential transcription from the uteroglobin promotor in endometrial cells. Nucleic Acids Res. **19**:2849–2859.
33. **Pilot-Matias, T. J., S. E. Kister, J. L. Fox, K. Kropp, S. W. Glasser, and J. A. Whitsett.** 1989. Structure and organization of the gene encoding human pulmonary surfactant proteolipid SP-B. DNA **8**:75–86.
34. **Poellinger, L., and R. G. Roeder.** 1989. Octamer transcription factors 1 and 2 each bind to two different functional elements in the immunoglobulin heavy-chain promoter. Mol. Cell. Biol. **9**:747–756.
35. **Ponce, E., J. A. Lloyd, A. Pierani, R. G. Roeder, and J. B. Lingrel.** 1991. Transcription factor OTF-1 interacts with two distinct DNA elements in the ^{A}gamma-globin gene promoter. Biochemistry **30**:2961–2967.
36. **Singh, G., and S. L. Katyal.** 1984. An immunological study of the secretory products of rat Clara cells. J. Histochem. Cytochem. **32**:49–54.
37. **Singh, G., J. Singh, S. L. Katyal, W. E. Brown, J. A. Kramps, I. L. Paradis, J. H. Dauber, T. A. MacPherson, and N. Squeglia.** 1988. Identification, cellular localization, isolation, and characterization of human Clara cell-specific 10 KD protein. J. Histochem. Cytochem. **36**:73–80.
38. **Solomon, M. J., F. Strauss, and A. Varshavsky.** 1986. A mammalian high mobility group protein recognizes any stretch of six AT base pairs in duplex DNA. Proc. Natl. Acad. Sci. USA **83**:1276–1280.
39. **Stripp, B. R., P. L. Sawaya, D. S. Luse, K. A. Wikenheiser, S. E. Wert, J. A. Huffman, D. L. Lattier, G. Singh, S. L. Katyal, and J. A. Whitsett.** 1992. Cis-acting elements that confer lung epithelial cell expression of the CC10 gene. J. Biol. Chem. **267**:14703–14712.
40. **Stripp, B. R., J. A. Whitsett, and D. L. Lattier.** 1990. Strategies for analysis of gene expression: pulmonary surfactant proteins. Am. J. Physiol. **259**:L185–L197.
41. **Suske, G., W. Lorenz, J. Klug, A. Gazdar, and M. Beato.** 1993. Elements of the rabbit uteroglobin promoter mediating its transcription in epithelial cells from the endometrium and lung. Gene Exp. **2**:339–351.
42. **Thanos, D., and T. Maniatis.** 1992. The high mobility group protein HMG I(Y) is required for NF-κB-dependent virus induction of the human IFN-β gene. Cell **71**:777–789.
43. **Tian, J. M., and U. Schibler.** 1991. Tissue-specific expression of the gene encoding hepatocyte nuclear factor-1 may involve hepatocyte nuclear factor-4. Genes Dev. **5**:2225–2234.
44. **Umland, T. C., S. Swaminathan, W. Furey, G. Singh, J. Pletcher, and M. Sax.** 1992. Refined structure of rat Clara cell 17 kDa protein at 3.0 Angstrom resolution. J. Mol. Biol. **224**:441–448.
45. **Wikenheiser, K. A., S. E. Wert, J. R. Wispé, M. Stahlman, M. D'amore-Bruno, G. Singh, S. L. Katyal, and J. A. Whitsett.** 1992. Distinct effects of oxygen on surfactant protein-B expression in bronchiolar and alveolar epithelium. Am. J. Physiol. **262**:L32–L39.
46. **Xanthopoulos, K. G., V. R. Prezioso, W. S. Chen, F. M. Sladek, R. Cortese, and J. E. Darnell, Jr.** 1991. The different tissue transcription patterns of genes for HNF-1, C/EBP, HNF-3, and HNF-4, protein factors that govern liver-specific transcription. Proc. Natl. Acad. Sci. USA **88**:3807–3811.