MOLECULAR AND CELLULAR BIOLOGY, Sept. 1994, p. 5671–5681
0270-7306/94/$04.00+0

Vol. 14, No. 9

The Lung-Specific Surfactant Protein B Gene Promoter Is a Target for Thyroid Transcription Factor 1 and Hepatocyte Nuclear Factor 3, Indicating Common Factors for Organ-Specific Gene Expression along the Foregut Axis

ROBERT J. BOHINSKI,[1] ROBERTO DI LAURO,[2] AND JEFFREY A. WHITSETT[1]*

Division of Pulmonary Biology, Children's Hospital Medical Center, Cincinnati, Ohio 45229-2899,[1] and Stazione Zoologica 'Anton Dohrn,' 80121 Naples, Italy[2]

Received 14 April 1994/Returned for modification 9 May 1994/Accepted 2 June 1994

We used the lung epithelial cell-specific surfactant protein B (SPB) gene promoter as a model with which to investigate mechanisms involved in transcriptional control of lung-specific genes. In a previous study, we showed that the SPB promoter specifically activated expression of a linked reporter gene in the continuous H441 lung cell line and that H441 nuclear proteins specifically protected a region of this promoter from bp −111 to −73. In this study, we further show that this region is a complex binding site for thyroid transcription factor 1 (TTF-1) and hepatocyte nuclear factor 3 (HNF-3). Whereas TTF-1 bound two highly degenerate and closely spaced sites, HNF-3 proteins bound a TGT3 motif (TGTTTGT) that is also found in several liver-specific gene regulatory regions, where it appears to be a weak affinity site for HNF-3. Point mutations of these binding sites eliminated factor binding and resulted in significant decreases in transfected SPB promoter activity. In addition, we developed a cotransfection assay and showed that a family of lung-specific gene promoters that included the SPB, SPC, SPA, and Clara cell secretory protein (CCSP) gene promoters were specifically activated by cotransfected TTF-1. We conclude that TTF-1 and HNF-3 are major activators of lung-specific genes and propose that these factors are involved in a general mechanism of lung-specific gene transcription. Importantly, these data also show that common factors are involved in organ-specific gene expression along the mammalian foregut axis.

The lung forms as an endodermal bud from the anteromedian foregut wall. Mesenchymal tissue interacts with this bud, inducing a process of branching morphogenesis that establishes a highly branched network of epithelially lined airways. Cellular differentiation in the lung is complex, and several morphologically distinct cell types comprise the airway epithelium. By definition, the onset of cellular differentiation is signalled by expression of differentiated gene products; hence, one fruitful approach to understanding mechanisms of cellular differentiation in mammals has been to identify factors that control expression of genes that define the cellular phenotype (14, 22, 34, 41, 71). Lung-specific gene products include the surfactant proteins (surfactant protein A [SPA], SPB, SPC, and SPD) and Clara cell secretory protein (CCSP). The recent cloning of these genes, the determination of their expression patterns in vivo (70, 72, 75), and the characterization of cell lines that support their expression (29, 55, 76) now provide a model system with which to investigate mechanisms involved in the establishment of several distinct cell types within a single endodermally derived organ. Some of these gene products overlap in expression and are expressed in at least two morphologically and spatially distinct cell types. The challenge will be to determine the mechanisms that permit both colocalized and segregated expression of these genes along the airway axis.

The control of tissue-specific gene expression occurs largely at the level of transcription initiation (21). Consistent with this observation is that appropriate *cis*-active sequences from tissue-specific genes are often sufficient to target expression of a reporter gene to the appropriate tissue in vivo (36). Extensive study has shown that DNA-binding proteins specifically interact with these sequences to stimulate gene transcription (37, 46, 50). One model suggests that the mechanism by which these proteins act depends on their restricted cellular distribution and interaction with only one type of tissue-specific gene family. The control of skeletal muscle cell differentiation by the MyoD family of transcription factor proteins supports this model. The expression of these proteins is restricted to the skeletal muscle cell lineage, and ectopic expression in vitro or in vivo is sufficient to convert most cell types to skeletal muscle-like cells (22, 49, 73, 74). This property of myogenic conversion due to a dominant-acting master regulatory gene is so far unique to the skeletal muscle lineage, since a similar phenomenon has not been observed for other cellular differentiation programs.

Liver-specific *cis*-active elements have been studied extensively, and several transcription factors, including hepatocyte nuclear factor 1 (HNF-1), HNF-3, HNF-4, C/EBP, and D-element-binding protein (DBP) (reviewed in reference 23), bind these regions and appear to act together to regulate transcription of liver-specific genes (17). In contrast to the MyoD model, none of these proteins appears to be restricted to liver cells (78). This finding suggests that mechanisms other than restricted expression of a transcription factor to a single cell type are responsible for the liver-specific activity of these genetic elements. This could involve combinatorial interactions between DNA-bound factors at a unique *cis*-active environment (48, 52) or between a DNA-bound factor and a non-DNA-bound cofactor (43, 46) or covalent modification of a factor that could result in changes in DNA site preference

* Corresponding author. Mailing address: Division of Pulmonary Biology, Children's Hospital Medical Center, Cincinnati, OH 45229-2899. Phone: (513) 559-7665. Fax: (513) 559-7868.

(38) or in affinities for other proteins. Thus, higher-order protein-protein and protein-DNA interactions are likely important determinants of specificity in these systems.

It has recently become apparent that mechanisms to achieve tissue-specific gene expression in the liver and lung might be very similar. This is primarily suggested by the robust expression of HNF-3 and C/EBP liver transcription factor family members in the lung (10, 41, 78) and by the recent finding that HNF-3 proteins bind a region of the CCSP gene promoter in vitro (3, 65). SPB is a particularly attractive model gene for investigating mechanisms involved in transcriptional control of lung-specific genes. In part, this is due to the recent availability of continuous lung cell lines (human NCI-H441-4 [H441] and mouse MLE-15) that express high levels of endogenous SPB mRNA (29, 55, 76). These cell lines allow detailed mapping of transcriptionally active genetic elements by transient transfection and provide an enriched source of relevant nuclear proteins for biochemical studies. This approach is advantageous given the complex heterogeneity of cell types within the lung. In addition, SPB is expressed in both proximal and distal lung epithelial cells (77) and may thus provide insight into mechanisms that distinctly control gene expression along the airway axis.

In a previous study, we determined that a proximal-promoter fragment (bp −218 to 41) of the SPB gene was sufficient for high-level expression of a linked reporter gene in H441 cells but not heterologous cell lines (5). We detected five discrete but closely spaced DNase I footprints in this region between bp −107 and 33. Two of these footprints were detected by using H441 but not HeLa cell nuclear extracts (5). In this study, we describe the functional interaction of thyroid transcription factor 1 (TTF-1) and HNF-3 with this region of the SPB promoter, and we present data suggesting that these factors are involved in a general mechanism of lung-specific gene transcription. Importantly, these data also show that common factors are involved in organ-specific gene transcription for foregut derivatives. In particular, the in vivo expression patterns of factors involved in liver-, lung-, and thyroid-specific gene transcription suggest that a particular combination of these factors may reflect an axial coding strategy for organ-specific gene expression along the foregut axis.

MATERIALS AND METHODS

Nuclear extract preparation. H441 and MLE-15 nuclear extracts were prepared by using a modified mini-extract procedure (66). All procedures for nuclear extraction were performed on ice and with ice-cold reagents. Confluent monolayers from one to four, 10-cm-diameter dishes were washed twice with 10 ml of ice-cold phosphate-buffered saline (PBS), harvested by scraping into 1 ml PBS, and pelleted in a 1.5-ml microcentrifuge tube at 3,000 rpm for 5 min. The cell pellet was washed once in 1 ml of PBS and pelleted as described above. The pellet was resuspended in 1 packed cell volume of fresh buffer A (10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; pH 7.9], 10 mM KCl, 0.1 mM EDTA, 1.5 mM $MgCl_2$, 0.2% [vol/vol] Nonidet P-40, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride), and cells were lysed during a 5-min incubation with occasional gentle vortexing. A nuclear pellet was obtained by microcentrifugation at 3,000 rpm for 5 min, and the supernatant was the cytoplasmic extract. The nuclear pellet was resuspended in 1 packed nuclear volume of fresh buffer B (20 mM HEPES [pH 7.9], 420 mM NaCl, 0.1 mM EDTA, 1.5 mM $MgCl_2$, 25% [vol/vol] glycerol, 1 mM dithrothreitol, 0.5 mM phenylmethylsulfonyl fluoride), and nuclei were extracted during a 10-min incubation with occasional gentle vortexing. Extracted nuclei were pelleted in a microcentrifuge at 14,000 rpm for 10 min. The supernatant was recovered and typically contained 5.0 to 10.0 μg of extracted nuclear protein per μl. Nuclear extracts were stored at −80°C without loss of activity for at least 6 months.

Synthetic oligonucleotides. Synthetic oligonucleotides were annealed at 10 μM in 100 μl of buffer M (10 mM Tris [pH 7.5], 10 mM $MgCl_2$, 50 mM NaCl) by placing the mixture in a preheated 95°C dry block, which was then slowly cooled to room temperature. The A_{260} was determined, and dilutions of this mixture were made in TE (10 mM Tris [pH 8.0], 1 mM EDTA) and used directly in an electrophoretic mobility shift assay (EMSA) as unlabeled competitor DNA. For use as the probe in the EMSA, 20 μl of the annealed mixture was gel purified by using a 4% BIOGEL and MERmaid kit as specified by the manufacturer (Bio 101). The A_{260} was determined, and 1.5 pmol of annealed and gel-purified oligonucleotide were end labeled by using [γ-^{32}P]ATP and T4 polynucleotide kinase. End-labeled probe was purified from unincorporated [γ-^{32}P]ATP by using a Pharmacia Nick Column and recovered in 400 μl of TE for an activity of approximately 25,000 dpm μl^{-1}.

EMSA. The EMSA was adapted from a previous protocol (33). Briefly, nuclear extract (1 to 2 μl) and, when indicated, unlabeled oligonucleotide competitor DNA were preincubated in 20 μl of buffer C [12 mM HEPES (pH 7.9), 4 mM Tris-Cl (pH 7.9), 25 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, 1 mM dithiothreitol, 50 ng of poly(dI-dC) (Boehringer Mannheim) μl^{-1}, 0.2 mM fresh phenylmethylsulfonyl fluoride] for 10 min on ice. Probe (100,000 dpm) was added, and the mixture was incubated for an additional 20 min on ice. For antibody supershift and interference assays, 1 μl of antibody was added after the addition of probe, and the mixture was incubated for an additional 20 min on ice. TTF-1 antibody is described by Lazzaro et al. (42). HNF-3α, -β, and -γ antibodies were kindly provided by J. E. Darnell, Jr. (41). Recombinant, bacterially expressed TTF-1 homeodomain protein (TTF-1 HD) is described by Damante and Di Lauro (18). Assays were performed with 1 μl of TTF-1 HD in place of nuclear extract. Bound and free probe were resolved by nondenaturing polyacrylamide gel electrophoresis. Five percent gels (acrylamide/bisacrylamide, 29:1; 0.5× TBE [44.5 mM Tris, 44.5 mM borate, 1 mM EDTA, pH 8.3]; 2.5% [vol/vol] glycerol; 1.5 mm thick) were run in 0.5× TBE running buffer at constant current (30 mA) for approximately 90 min. Gels were blotted to Whatman 3MM paper, dried under vacuum, and exposed to X-ray film for 1 to 3 h at −80°C with an intensifying screen.

Plasmid constructions and site-directed mutagenesis. The human SPB gene promoter (bp −218 to 41) was isolated from p2244/436 (5) by using PCR and linker primers to create 5′ *Hin*dIII and 3′ *Sal*I sites. The product was digested with *Hin*dIII and *Sal*I and cloned into the respective sites of M13mp18 and used as the template for site-directed mutagenesis performed by the method of Kunkel (40). The wild-type and mutated promoters were isolated from M13 replicative form by *Hin*dIII and *Sal*I digestion and cloned into the respective sites of pBLCAT6 (8). These SPB promoter-chloramphenicol acetyltransferase (CAT) fusion plasmids were designated p218/41-WT, -5T, -3T, -TT, -H, and -TTH, and identities were confirmed by dideoxy sequencing of double-stranded templates. The 5′ deletion mutant pΔ-80 contains the human SPB gene promoter (bp −80 to 41) in the *Hin*dIII and *Sal*I sites of pBLCAT6 and was made by using PCR and linker primers as described above. The rat CCSP gene promoter (bp −2338 to 49) was cloned into the polylinker of pBLCAT6 as

described previously (70) and was kindly provided by B. R. Stripp. The mouse SPC gene promoter (bp −4680 to 18), a gift from S. Glasser (Children's Hospital, Cincinnati, Ohio), was isolated as an *Xba*I and *Hpa*II fragment, digested with nuclease *Bal* 31 at its 3′ end, repaired with T4 DNA polymerase to bp 18, and cloned as an *Xba*I- and 3′ *Xho*I-linked fragment into the respective sites of pBLCAT6. pBLCAT5 contains the thymidine kinase gene promoter (bp −105 to 51) (8). TTR-CAT contains the mouse transthyretin (TTR) gene promoter (bp −202 to 9) and was kindly provided by J. E. Darnell, Jr. (41).

Cell culture, transfections, and reporter gene assays. H441 and MLE-15 cells (used in the nuclear extract procedure) were maintained exactly as described previously (55, 76). HeLa cells were maintained in Dulbecco's modified Eagle medium containing 10% heat-inactivated fetal bovine serum. The day before transfection, confluent monolayers were split (1:5 to 1:8 for H441 cells and 1:20 for HeLa cells) into 10-cm-diameter dishes. Four hours before transfection, cells were switched to transfection medium (Dulbecco's modified Eagle medium containing 10% heat-inactivated fetal bovine serum and 1% penicillin-streptomycin; Gibco BRL). Transfections were performed by the calcium phosphate coprecipitation method as described by Rosenthal (62) except that glycerol shock was not used. For the analysis of point mutants in H441 cells, precipitates were prepared by using 5.0 pmol of promoter-CAT fusion plasmid and 2.5 pmol the internal control plasmid, pCMV-βgal (44), per 10-cm-diameter dish. For TTF-1 transactivation experiments in HeLa cells, each 10-cm dish was treated with a precipitate prepared by using 15.0 μg of promoter-CAT fusion plasmid, 2.0 μg of pCMV-βgal, 7.5 μg of pUC19, and 0.5 μg of either the empty expression vector (pRc/CMV; Invitrogen) or an expression vector containing the entire TTF-1 open reading frame (pCMV-TTF-1 [28]). Precipitates were added dropwise to the medium covering the cells. Cells were incubated with precipitate for 14 to 18 h, washed once with calcium- and magnesium-free Hanks' balanced salt solution, returned to maintenance medium, and cultured for an additional 24 h (H441 cells) or 48 h (HeLa cells). Cells were harvested, freeze-thaw lysates were prepared in 100 μl of 0.25 M Tris (pH 7.8), and aliquots were assayed for CAT and β-galactosidase activities essentially as described previously (45, 62). To correct for variations in transfection efficiency, lysates were normalized for β-galactosidase activity so that CAT enzyme assays contained equivalent amounts of β-galactosidase activity. Thin-layer chromatograms of [^{14}C] chloramphenicol and its acetylated derivatives were quantitated with a Molecular Dynamics PhosphorImager.

RESULTS

Multiple and conserved nuclear factors bind the SPB gene promoter. In previous work, we identified a region of the human SPB gene promoter specifically protected by lung cell nuclear proteins in DNase I footprinting experiments (5). Comparison with homologous sequences from the mouse SPB gene promoter revealed two 14-bp blocks of uninterrupted identity within these footprinted regions (Fig. 1A). We used this 55-bp region as a probe in EMSA and observed several specific and nonspecific complexes (6). We simplified resolution of these complexes and reduced nonspecific binding by designing subprobes of this region based on the blocks of conserved sequence and DNase I protection. This resulted in two probes, designated SPB-f1 and SPB-f2 (Fig. 1A). To aid identification of important complexes, we exploited the evolutionary conservation of this region and the idea that the cognate cell-type-specific transcription factors would also be conserved. We used human (H441) and mouse (MLE-15) lung adenocarcinoma cell lines, which express the endogenous SPB gene, and DNA binding site homologs for comparison. Nuclear extracts from both cell lines formed two complexes of identical electrophoretic mobility with SPB-f1 (Fig. 1B, lanes 1 and 2, arrows A and B) and, similarly, one complex of identical electrophoretic mobility with SPB-f2 (Fig. 1B, lanes 3 and 4, arrow C). Complex D (Fig. 1B, lane 3, arrow D) resolved from complex C by extended electrophoresis and only appeared in MLE-15 nuclear extracts. A complex of low abundance and high mobility, apparent with H441 nuclear extract and SPB-f1 (Fig. 1B, lane 2), was not reproducible under these conditions. 3T3 nuclear extracts contained AP1 binding activity but did not retard SPB-f1 or SPB-f2, and HeLa nuclear extracts contained only a small amount of complex C (6). Because we wanted to identify complex D as well as conserved complexes A to C, we used MLE-15 nuclear extract for further study. We determined the binding specificity of these complexes by addition of unlabeled competitor oligonucleotides. This resulted in efficient competition for complexes A and B or C and D by an excess of self (Fig. 1C and D, lanes 2 and 3) the mouse homolog of self (Fig. 1C and D, lanes 4 and 5) but not the respective adjacent binding site (Fig. 1C and D, lanes 6 and 7). Although SPB-f1 and SPB-f2 competed for complex A, SPB-f1 was clearly a more efficient competitor (Fig. 1C, lanes 2 and 6). In addition, complex A was always eliminated in competition assays before complex B, suggesting that complex A might be a higher-order complex dependent on complex B formation. In contrast, the relative amounts of complexes C and D were not sensitive to protein concentration, suggesting that two different factors bound SPB-f2 in a mutually exclusive way. For SPB-f2, the human sequence appeared to be a better competitor than the mouse homolog, but both were significantly more efficient competitors than the adjacent SPB-f1 binding site. Because SPB-f1 and SPB-f2 did not cross-compete in these assays, we concluded that at least two distinct and evolutionarily conserved nuclear factors specifically bound this region.

A TGT3 motif binds HNF-3 proteins expressed in the lung. SPB-f2 contained an obvious TGT3 motif (TGTTTGT) that occurs in the regulatory elements of diverse liver-specific genes (reference 35 and references therein). Because of its apparent novelty, this motif was also termed HNF-5 to distinguish it from motifs recognized by other liver transcription factors, including HNF-3 (30, 61). Surprisingly, this motif binds HNF-3 proteins (24, 35, 53, 57). As shown in Fig. 2A, SPB-f2 was not as clearly related to the prototypical HNF-3 motif present in the liver-specific TTR and α1-antitrypsin regulatory regions (17). We used oligonucleotides typical of each HNF-3 motif as unlabeled competitors in EMSA and found that both motifs efficiently competed for complexes C and D. A TGT3 site from the tyrosine aminotransferase gene enhancer (30) or the strong HNF-3 site from the TTR gene promoter (TTR-S [17]) were efficient competitors for complexes C and D (Fig. 2B), but a mutant TGT3 motif (mTGT3) that does not bind HNF-3 (30) did not compete for complex C or D (Fig. 2B). We added to EMSA reaction mixtures antisera specific for each HNF-3 protein (anti-HNF-3α, -β, and -γ [41]) and showed the binding of both HNF-3α and HNF-3β to SPB-f2 in MLE-15 nuclear extracts. Anti-HNF-3α and -β significantly interfered with the formation of complexes C and D, respectively, and formed only minor supershifted complexes of lower mobility (Fig. 2C, α* and β*). The identification of the lowest-mobility complex as HNF-3α was consistent with the relative mobilities of HNF-3 proteins in liver cells, in which HNF-3β complexes migrate only slightly faster than HNF-3α complexes and the two

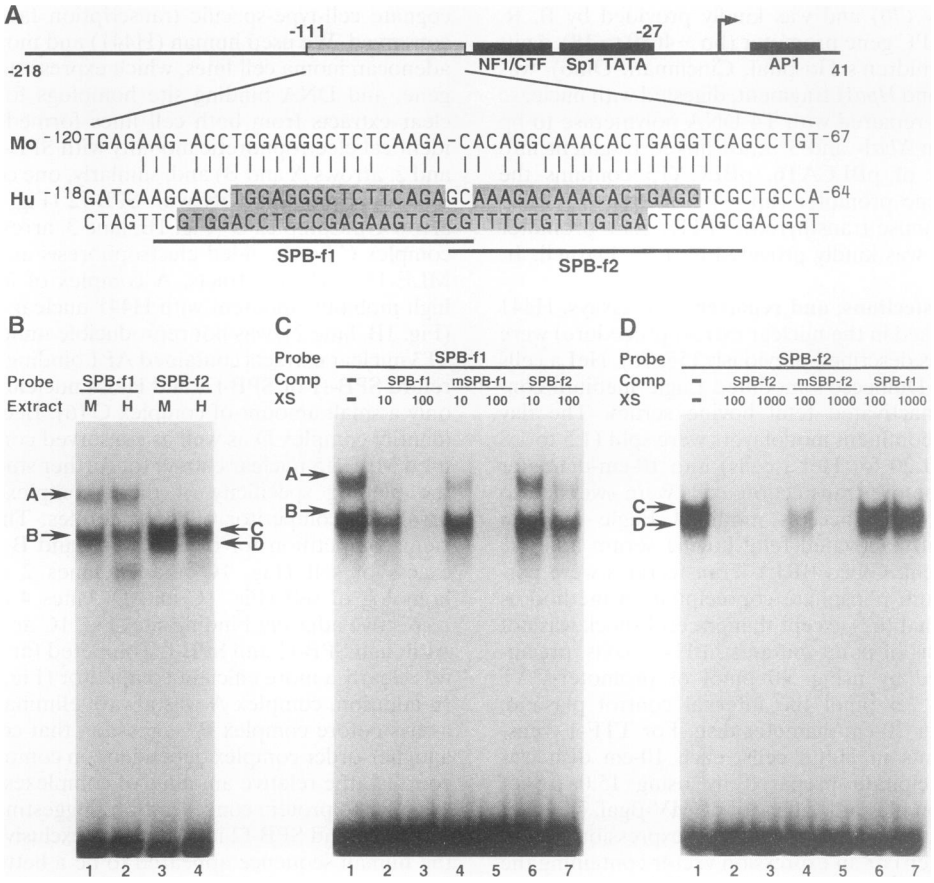

FIG. 1. Evolutionary conservation of nuclear factors bound to the SPB gene promoter. (A) Schematic illustration of the SPB gene promoter region (bp −218 to 41) showing the locations of DNase I footprints identified in our previous study (5). This promoter fragment is active in H441 cells but not A549 or HeLa cells (5). Shaded rectangles are lung-enriched footprints, and black rectangles are ubiquitous footprints. The transcription start site is indicated by an arrow. Ubiquitous footprints are labeled to indicate putative *cis*-active binding sites for ubiquitous nuclear factors NF1/CTF, Sp1, and AP1. Homologous mouse (Mo) and human (Hu) nucleotide sequences that span the region protected by lung-enriched nuclear proteins are aligned below. Vertical lines indicate identity between the mouse and human sequences, and dashes are gaps inserted for maximal alignment. Shaded nucleotides in the human sequence indicate the precise pattern of DNase I protection. Thick horizontal lines denote the EMSA probes, SPB-f1 and SPB-f2, derived from this sequence. (B) Comparison of MLE-15 (M) and H441 (H) nuclear extracts in EMSA with SPB-f1 and SPB-f2 probes. (C and D) Specific binding of distinct nuclear proteins to SPB-f1 and SPB-f2. Unlabeled competitors (Comp) were added to EMSA mixtures at the fold molar excess (XS) indicated above each lane. mSPB-f1 and mSPB-f2 are the mouse oligonucleotide homologs to each human probe. Complexes A to D are indicated to the left or right of each panel by arrows.

complexes appear as a single broad band in EMSA (41). Simultaneous addition of both anti-HNF-3α and anti-HNF-3β eliminated all major complex formation with SPB-f2, indicating that other proteins did not independently bind this region (Fig. 2C, lane 5). Anti-HNF-3α and -β did not significantly affect complexes A and B (Fig. 2C, lanes 6 to 8), again supporting the idea that factors bound to SPB-f1 were distinct. These reactions were specific because an identical antibody preparation against HNF-3γ did not specifically modify major complex formation (Fig. 2C, lane 4), consistent with the lack of HNF-3γ expression in the lung (41). H441 nuclear proteins formed only complex C, and we determined this was HNF-3α (6). In addition, we used Northern (RNA) blot analysis and detected transcripts for HNF-3α and -β in MLE-15 cells and only HNF-3α in H441 cells (67).

Highly degenerate motifs bind TTF-1 in close apposition. An obvious consensus *cis*-active motif was not apparent in SPB-f1. In EMSA, complex A appeared at high nuclear protein concentration and was eliminated before complex B by unlabeled self competitor (6). This result suggested that two factors might independently bind SPB-f1 to create a trimeric protein-DNA complex. We tested this hypothesis by using 5′ (5′f1) or 3′ (3′f1) subfragments of SPB-f1 as competitors and probes in EMSA (Fig. 3). Subfragments 5′f1 and 3′f1 were specific and equivalent competitors for complexes A and B (Fig. 3B, lanes 1 to 5). When labeled and used as probes, the two subfragments formed a single complex of identical mobility (Fig. 3B, lanes 6 and 7). This result suggested that the same factor bound each end of SPB-f1 and prompted a detailed comparison of these sequences. Maximal alignment of subfragments 5′f1 and 3′f1 showed less than 50% identity. However, the consensus from this alignment suggested a possible core *cis*-active motif (GCNCTNNAG) that greatly facilitated comparison to a list of *cis*-active motifs for vertebrate-encoded transcription factors (25). We found the limited identity with the TTF-1 binding site consensus (CCACTCAAGT) most attractive because this factor is expressed in developing lung epithelium (42) and is known to regulate thyroid-specific gene expression (27). We therefore considered the possibility that TTF-1 binds these highly degenerate sites in the SPB gene

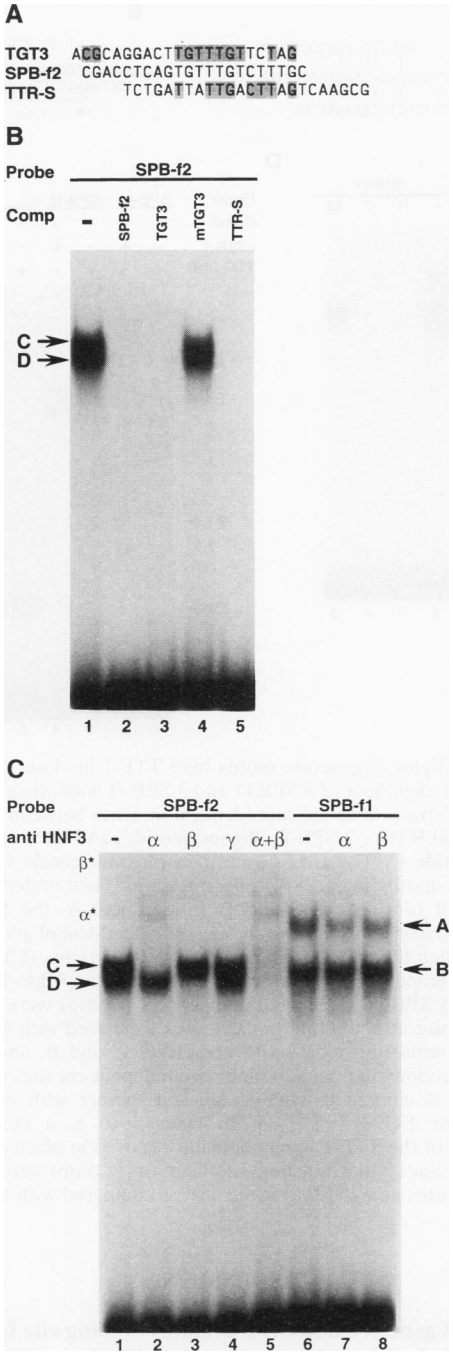

FIG. 2. TGT3 motif binds HNF-3 proteins. (A) Comparison of SPB-f2 with TGT3 and TTR-S HNF-3 binding sites. Nucleotides that match SPB-f2 are shaded in the TGT3 and TTR-S oligonucleotides. TGT3 is oligonucleotide s4 (30) from the tyrosine aminotransferase gene enhancer. TTR-S is oligonucleotide TTR-S (17) from the TTR gene promoter. mTGT3 contains a 2-bp mutation that eliminates specific binding of HNF-3 and is the same as oligonucleotide s4 mut (30). (B) HNF-3 binding sites are efficient competitors for complexes C and D. Unlabeled competitors (Comp) were added to EMSA mixtures at a 1,000-fold molar excess compared with the probe. (C) Antisera to HNF-3α and -β react with MLE-15 nuclear proteins bound to SPB-f2. Antiserum to each HNF-3 protein was added to EMSA reaction mixtures with MLE-15 nuclear extracts as indicated above each lane. In this assay, HNF-3 antisera (α and β) primarily interfered with complex formation. Asterisks indicate the positions of weak supershifted complexes that were detected by extended autoradiographic exposure. A weak band was formed by all antisera added to this assay and was thus nonspecific.

promoter. Using the same strategy as used for the identification of HNF-3 (see above), we found that a high-affinity binding site for TTF-1 from the thyroid-specific thyroglobulin gene promoter (oligonucleotide C [12]) was a more efficient competitor for complex A and B than self (Fig. 4C, lanes 2 to 4). When labeled and used as a probe, oligonucleotide C formed a complex of mobility identical to that of complex B, consistent with the single TTF-1 binding site in this oligonucleotide (Fig. 4C, lanes 1 and 2). We used affinity-purified, polyclonal antisera to TTF-1 in EMSA reactions (anti-TTF-1 [42]) and showed binding of TTF-1 at two independent sites in SPB-f1. Addition of anti-TTF-1 to EMSA reaction mixtures containing either MLE-15 or H441 nuclear proteins resulted in the elimination of complexes A and B and the formation of a lower-mobility complex of similar abundance (Fig. 4D, lanes 1 and 2) (6). This reaction was specific because anti-TTF-1 did not alter HNF-3 and SPB-f2 complex formation (Fig. 4D, lanes 3 and 4). Further, a recombinant fragment of TTF-1 containing only the homeodomain (TTF-1 HD [18]) specifically bound SPB-f1 and formed a two-banded pattern (A′ and B′ in Fig. 4D, lanes 5 to 8). Complex A′ formed at higher protein concentrations and depended on the integrity of both sites in SPB-f1 (6) (Fig. 5C). Mutation of bases within either site resulted in complete loss of complex A′ and a reduction in complex B′ (Fig. 5C, lanes 2 and 3). Mutation of both sites completely eliminated specific binding of TTF-1 HD (Fig. 5C, lane 4). This result indicated that complex B was TTF-1 bound at a single site while complex A was TTF-1 bound at two independent sites in oligonucleotide SPB-f1. In addition, we detected robust expression of TTF-1 transcripts in both MLE-15 and H441 cells (67). We conclude that TTF-1 is present in both MLE-15 and H441 cells and specifically binds at least two highly degenerate motifs in the SPB promoter.

TTF-1 and HNF-3 protein binding sites are critical for SPB gene promoter activity. Figure 5A summarizes the relative locations of TTF-1 and HNF-3 binding sites identified in this study. We next constructed mutations at each site and showed that binding was dependent on a specific sequence because a 2-bp mutation at each site severely impaired factor binding in EMSA experiments. The mutated version of SPB-f2 (H) did not compete for or bind HNF-3 proteins (Fig. 5D), and as discussed above, TTF-1 binding depended on bases within the consensus motif (Fig. 5C). To determine if these sites were transcriptionally active, we used site-directed mutagenesis to construct these binding site mutations in the SPB gene promoter. We linked the wild-type (WT) and mutant promoters to a CAT reporter gene and assayed for transcriptional activity in H441 and HeLa cells (Fig. 5B). All mutations resulted in a statistically significant reduction in CAT activity in H441 cells, and no mutation affected activity in HeLa cells, thus demonstrating the restricted cellular activity of factors bound to this region. Mutation of the 5′ TTF-1 binding site (5T) was less dramatic than the 3′ TTF-1 binding site (3T), and mutation of both TTF-1 sites (TT) was no different than for the 3T mutation, suggesting that the 5′ site depended on the 3′ site for activity. Mutation of all three binding sites (TTH) resulted in an activity that was not different from gross deletion of all sequences upstream of −80 (Δ−80). This indicated that either no other sites were present between −218 and −80 or that no other site in this region could effect SPB promoter activity in

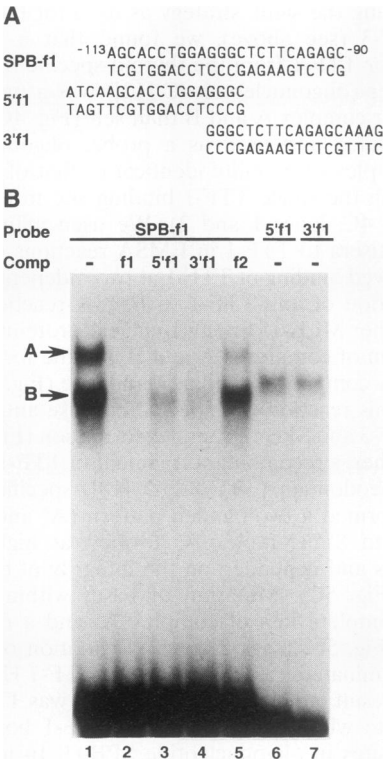

FIG. 3. SPB-f1 contains two similar binding sites. (A) Subfragments of SPB-f1 are identical competitors in EMSA. Subfragments were extended 4 bp beyond SPB-f1 in this region to prevent the oligonucleotide from being too small for EMSA. (B) Unlabeled competitors (Comp) were added to EMSA mixture at a 100-fold molar excess compared with the probe.

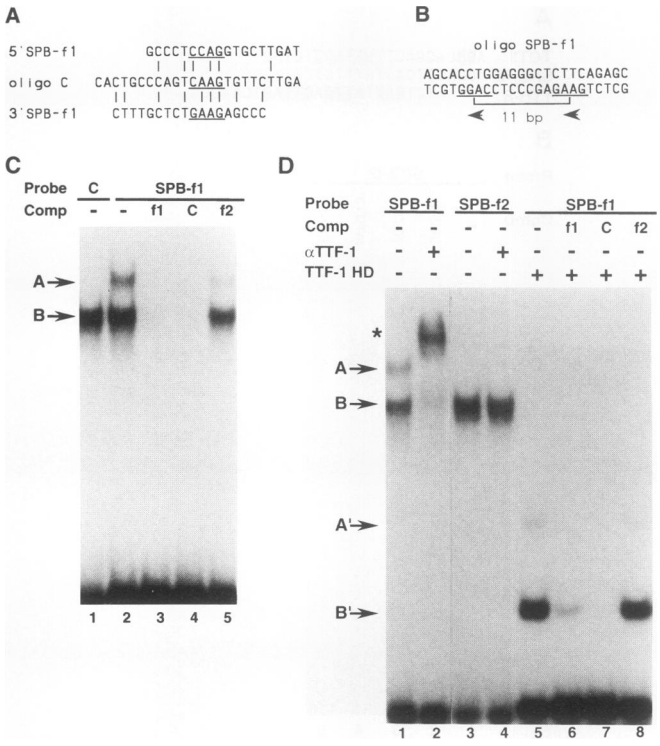

FIG. 4. Highly degenerate motifs bind TTF-1 in close apposition. (A) Optimal alignment of 5'SPB-f1 and 3'SPB-f1 with oligonucleotide (oligo) C. Vertical lines indicate identical matches between the lower strand of 5'SPB-f1 or 3'SPB-f1 oligonucleotides and the top strand of oligonucleotide C. The CAAG motif in oligonucleotide C and the homologous motifs in the SPB oligonucleotides are underlined. (B) Arrangement of CAAG-like motifs (underlined) in the SPB gene promoter. Arrowheads indicate the possible head-to-tail arrangement of TTF-1 when bound to both sites separated by 11 bp. (C) Oligonucleotide C competes for complex formation between MLE-15 nuclear proteins and SPB-f1. Unlabeled competitors (Comp) were added to EMSA mixture at a 100-fold molar excess compared with the probe. (D) TTF-1 antiserum reacts with complexes A and B, and purified TTF-1 homeodomain fragment binds two independent sites in SPB-f1. Lanes 1 to 4 contained MLE-15 nuclear extract with or without antiserum to TTF-1 (αTTF-1). In lanes 5 to 8, a recombinant preparation of the TTF-1 homeodomain was used in place of nuclear extract. For lanes 5 to 8, unlabeled competitors (Comp) were added to EMSA mixtures at a 100-fold molar excess compared with the probe.

the absence of the TTF-1 and HNF-3 sites that we identified. We conclude that these TTF-1 and HNF-3 protein binding sites are critical to SPB gene promoter activity in H441 cells.

TTF-1 is a strong *trans* activator of the SPB gene promoter and other lung-specific gene promoters. We reasoned that TTF-1 would function as a binding site-dependent *trans* activator of the SPB gene promoter and other target promoters and used the SPB gene promoter and binding site mutants to develop a cellular assay for DNA binding and transcriptional activation by TTF-1. We cotransfected a TTF-1-deficient cell line (HeLa) with WT or mutant SPB gene promoters and either an empty expression vector or one carrying the full-length TTF-1 cDNA (28). TTF-1 dramatically increased activity from the WT SPB gene promoter (Fig. 6A, lanes 1 and 2) and the HNF-3 mutant promoter (Fig. 6A, lanes 5 and 6) but had no effect on the TTF-1 mutant promoter (Fig. 6A, lanes 3 and 4). Because TTF-1 transactivation depended on the integrity of TTF-1 binding sites, these results further demonstrated a direct effect of TTF-1 on SPB gene promoter activity. We exploited this system and determined the transcriptional responses of other lung-specific promoters to TTF-1. TTF-1 dramatically increased the activity of the lung-specific CCSP and SPC gene promoters but had no effect on the liver-specific TTR or the constitutive thymidine kinase gene promoter (Fig. 6B). In separate but identical experiments, the lung-specific SPA gene promoter also dramatically responded to cotransfected TTF-1 (6).

DISCUSSION

The SPB gene promoter is a complex binding site for TTF-1 and HNF-3. In this study, we showed the binding of TTF-1 and HNF-3 to closely spaced sites in the SPB gene promoter. The identification of TTF-1 binding sites was greatly facilitated by the observation that two adjacent sites appeared to bind the same factor. This allowed determination of a minimal consensus binding site motif (GCNCTNNAG) that partially matched the consensus TTF-1 binding site compiled from thyroid-specific genes (CCACTCAAGT [27]). It is interesting that these motifs contain a common element of dyad symmetry (CTNNAG). The consensus binding site for TTF-1 compiled from thyroid-specific genes contains the dyad, ACTCAAGT, and the two CTNNAG dyads found in the SPB gene promoter are part of larger but distinct dyads, CCTGGAGG and GCTCTTCAGAGC. TTF-1 is a member of the homeodomain class of DNA-binding proteins (31). Most homeodomain-

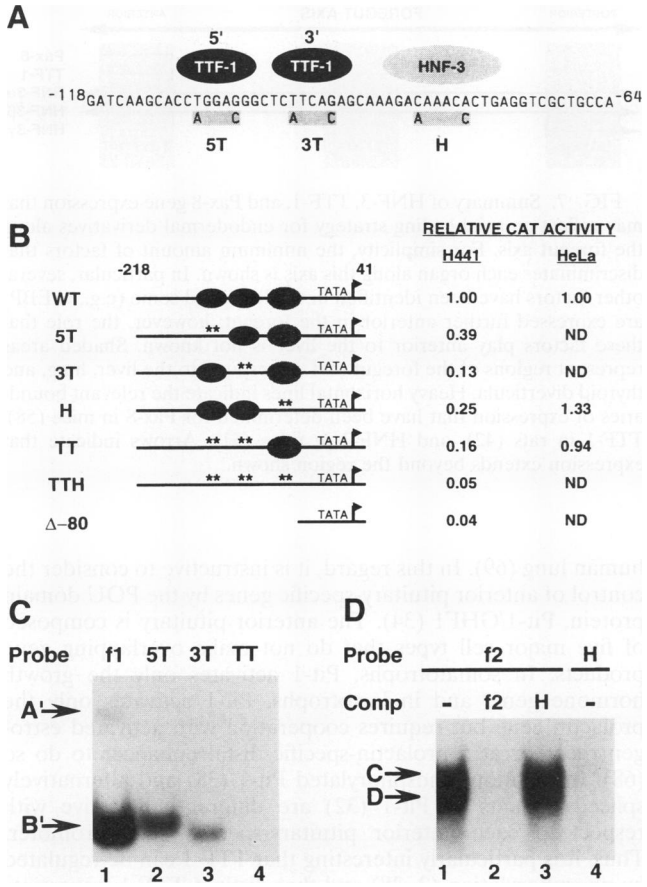

FIG. 5. TTF-1 and HNF-3 binding sites are critical to SPB gene promoter activity in H441 cells. (A) The locations of TTF-1 and HNF-3 binding sites identified in the SPB gene promoter are shown at the top. The shaded nucleotides below the sequence indicate the 2-bp mutations that were made at each binding site. (B) Mutation of TTF-1 and HNF-3 binding sites significantly reduces SPB gene promoter activity in H441 cells. The WT and mutant constructs are schematically shown. Constructs were transiently transfected into the H441 and HeLa cell lines with the pCMVβ-gal internal control plasmid. CAT activity was normalized to β-galactosidase activity, and the CAT activity of each mutant construct was then compared with the activity of the WT construct (assigned a value of 1) in each cell line. ND, not determined. The results shown are average values from three independent experiments in which the standard error of the mean was less than 10%. (C and D) Mutation of HNF-3 and TTF-1 binding sites eliminates factor binding in vitro. EMSA reactions were performed on mutant probes that were the same as wild-type SPB-f1 and SPB-f2 but contained the 2-bp mutations shown in panel A. Only the portion of the gel containing bound complexes is shown. In panel C, 1 μl of TTF-1 HD was used in place of nuclear extract and incubated with the WT SPB-f1 probe (f1) or the mutant probes (5T, 3T, and TT) in EMSA. In panel D, the WT SPB-f2 probe was compared with mutant probe H in EMSA using MLE-15 nuclear extract. Unlabeled competitors (Comp) were added to EMSA mixtures at a 1,000-fold molar excess compared with the probe.

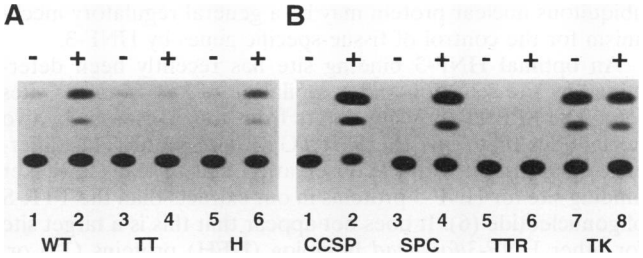

FIG. 6. Transactivation of lung-specific promoters by TTF-1 in HeLa cells. (A) The WT, TT, or H SPB promoter construct was transiently cotransfected with the internal control plasmid pCMVβ-gal and either an empty expression vector (−) or an expression vector containing the full-length TTF-1 cDNA (+) into the HeLa cell line. Each determination is representative of three independent experiments that were normalized for β-galactosidase activity. (B) CCSP, SPC, TTR, and thymidine kinase (TK) gene promoter constructs were cotransfected into the HeLa cell line as described above, and results are representative of three independent experiments.

containing proteins recognize their cognate binding motifs as monomers, and despite the apparent dyad symmetry of the TTF-1 binding sites discussed above, TTF-1 also binds DNA as a monomer (20, 31). Although the appearance of these dyads is compelling, it is not clear if they are important determinants of TTF-1 binding site recognition. This is primarily suggested by the observation that a very strong TTF-1 binding site, oligonucleotide C, does not contain a perfect CTNNAG motif (Fig. 4A). In addition, methylation interference studies of TTF-1 bound to oligonucleotide C emphasizes CAAGTG as an important core motif in which the TTF-1 homeodomain is directionally oriented with respect to the CAAG motif (underlined in Fig. 4A) and the following TG dinucleotide (19). This allows precise alignment of SPB TTF-1 binding sites with oligonucleotide C (Fig. 4A) and suggests that the central CAGG and GAAG motifs of the 5′ and 3′ SPB-f1 sites are homologs of the CAAG motif in oligonucleotide C (underlined in Fig. 4A). It is particularly noteworthy that this alignment offsets the alignment of the two SPB CTNNAG motifs, thus deemphasizing the importance of this motif. Interestingly, these CAAG-like core motifs are separated by 11 bp (approximately one helical turn of DNA) in the context of the natural SPB gene promoter (Fig. 4B) and may indicate a unique functional arrangement for TTF-1 binding sites, since this appears to place the two TTF-1 proteins bound to these sites on the same face of the DNA helix in close head-to-tail apposition.

Identification of the HNF-3 binding site was simplified by recognition of an apparently distinct subclass of HNF-3 binding site, the TGT3 motif. This motif was originally identified in promoter and enhancer regions of diverse liver-specific genes and often in close apposition to other transcription factors (reference 35 and references therein). A recent study (35) has shown that a TGT3 motif from the albumin enhancer (site eH) binds HNF-3 proteins in tight apposition to a site bound by ubiquitous nuclear factor 1/CCAAT transcription factor (NF1/CTF) proteins. The binding site juxtaposition of HNF-3 and NF1/CTF creates a composite regulatory element that is sensitive to the local sequence environment. This complex activates gene transcription in the context of the native albumin enhancer and during hepatocyte differentiation but inhibits an activation function of HNF-3 when isolated and placed upstream of a minimal TATA promoter. The family of NF1/CTF proteins bind to the sequence $TGGCN_7CCA$ (7, 63). At the albumin eH site, the TGT3 motif is 2 bp downstream of the NF1/CTF motif (TGT3-NN-$TGGCN_8CA$). Interestingly, in the SPB gene promoter, an NF1/CTF motif is 1 bp upstream from the TGT3 motif ($TGGCN_8CA$-N-TGT3), and this site is protected by a ubiquitous nuclear protein (5). This finding suggests that the binding site juxtaposition of HNF-3 and a

ubiquitous nuclear protein may be a general regulatory mechanism for the control of tissue-specific genes by HNF-3.

An optimal HNF-3 binding site has recently been determined by site selection and compilation of known target sites (56). The SPB HNF-3 site differs from this consensus by two nucleotides that flank the SPB TGT3 core motif. This difference is consistent with the observation that this was a weaker binding site for HNF-3 proteins in our extracts than the TTR-S oligonucleotide (6). It does not appear that this is a target site for other HNF-3/*forkhead* homolog (HFH) proteins (15) or, alternatively, that they are present in MLE-15 or H441 lung cell lines because all major complex formation with this site was eliminated by antibodies to both HNF-3α and HNF-3β. This is important since several other HFH proteins (e.g., HFH-1 and HFH-4) appear enriched in the lung, and HFH-1 can bind at least a subset of target sites that also bind HNF-3 proteins (15, 56). Although it is noteworthy that we detected both HNF-3α and HNF-3β in MLE-15 cells but only HNF-3α in H441 cells, the significance of this difference is not clear. This may reflect a species-specific difference since several other mouse lung epithelial cell lines that express surfactant protein genes (76) also express transcripts for both HNF-3α and HNF-3β (67), and these transcripts appear to be expressed in similar regions of the developing mouse lung epithelium (51).

Common factors for organ-specific gene expression along the foregut axis. TTF-1 and HNF-3 were originally identified as transcriptionally active DNA-binding proteins involved in regulation of thyroid- and liver-specific genes, respectively (12, 16, 17, 31, 41). Interestingly, our study showed that these factors are similarly involved in transcriptional regulation of the lung-specific SPB gene promoter. Taken together, these findings suggest that common factors are involved in organ-specific gene transcription along the foregut axis. Our finding is likely to be prototypic, suggesting similar functional roles for other transcription factors expressed in distinct endodermal derivatives. In particular, members of the C/EBP transcription factor family are expressed in both lung and liver and are known to regulate liver-specific genes late in cellular differentiation (1, 10). Although these proteins are major determinants of cell-type-specific promoter activity in thyroid, liver, and lung cell lines, additional levels of regulation must determine which TTF-1 and HNF-3 target genes are expressed in each cell type. Several recent studies suggest a critical role for context-dependent and combinatorial factor interactions in setting the specificity of a particular genetic element (43, 48, 52), and it is appealing to consider that such mechanisms may be prototypic for control of distinct tissue-specific genes by HNF-3 and TTF-1.

In this study, we also showed that cotransfected TTF-1 specifically increased the activity of other known lung-specific gene promoters, including the CCSP, SPC, and SPA gene promoters (Fig. 6B) (6). Although further studies are required to determine if this effect is the result of direct *trans* activation of these other lung-specific promoters by TTF-1, *trans* activation of the SPB gene promoter was direct since this response depended on the presence of multiple TTF-1 binding sites. Lung-specific genes overlap in expression, but for the most part CCSP is a marker for the more proximal nonciliated bronchiolar cell and SPC is a marker for the more distal alveolar type II cell (75), while SPA and SPB are often found in both cell types (58). This finding suggests that TTF-1 may control the expression of each lung-specific gene in mutually exclusive cellular environments. Consistent with this idea, we have detected TTF-1 protein expression in both alveolar type II cells and nonciliated bronchiolar cells in late-gestation human lung (69). In this regard, it is instructive to consider the control of anterior pituitary-specific genes by the POU domain protein, Pit-1/GHF1 (34). The anterior pituitary is composed of five major cell types that do not make overlapping gene products. In somatotrophs, Pit-1 activates only the growth hormone gene, and in lactotrophs, Pit-1 activates only the prolactin gene but requires cooperation with activated estrogen receptor at a prolactin-specific distal enhancer to do so (68). In addition, phosphorylated Pit-1 (38) and alternatively spliced variants of Pit-1 (32) are differentially active with respect to each anterior pituitary-specific gene promoter. Thus, it is particularly interesting that TTF-1 can be regulated by phosphorylation (2, 28) and that distinct TTF-1 transcripts appear to be present in lung but not thyroid cells (31).

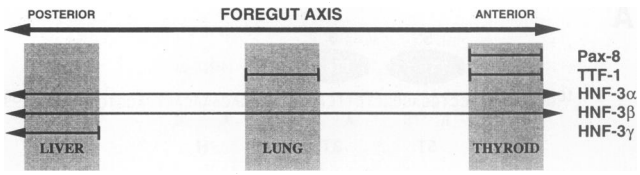

FIG. 7. Summary of HNF-3, TTF-1, and Pax-8 gene expression that may reflect an axial coding strategy for endodermal derivatives along the foregut axis. For simplicity, the minimum amount of factors that discriminates each organ along this axis is shown. In particular, several other factors have been identified in the liver, and some (e.g., C/EBP) are expressed further anterior in the foregut; however, the role that these factors play anterior to the liver is not known. Shaded areas represent regions of the foregut that correspond to the liver, lung, and thyroid diverticula. Heavy horizontal lines indicate the relevant boundaries of expression that have been determined for Pax-8 in mice (58), TTF-1 in rats (42), and HNF-3 in mice (51). Arrows indicate that expression extends beyond the region shown.

Regionalization of the foregut by HNF-3 and TTF-1. It is interesting to discuss our results in the context of what is known about the spatial and temporal expression and functional role of TTF-1 and HNF-3. TTF-1 is expressed in lung and thyroid epithelium and in restricted regions of the fetal brain (42). The lung and thyroid develop in close proximity as two endodermal diverticula from the anteromedian pharyngeal wall, and TTF-1 transcripts are restricted to the anterior migrating edge of the two diverticula (42). HNF-3 proteins comprise a gene family (α, β, and γ) and are expressed early in mouse development at the time of gastrulation and in all three germ layers (1, 51, 64), but expression is restricted to endoderm by embryonic day 15.5 (51). In particular, endodermal expression is dictated by the anterior boundary of each gene. HNF-3α and -β transcripts are detected in all endoderm-derived structures, while HNF-3γ has an anterior boundary in the liver (1, 51, 64). This distinguishes HNF-3 gene expression for the liver, which also develops as a diverticulum from the anteromedian foregut wall, from the more cranial lung and thyroid diverticula. Thus, coexpression of TTF-1 and HNF-3α and -β appears to define the cranial foregut, while coexpression of HNF-3α, -β, and -γ defines hepatic and posterior endodermal structures. Pax-8, a paired-domain protein, is expressed in the thyroid, central nervous system, and kidney (59). It binds in a mutually exclusive way to sites that overlap with TTF-1 and is a *trans* activator of thyroid-specific genes (79). The expression of Pax-8 in thyroid distinguishes this part of the foregut from all posterior endodermal structures, including the lung. Taken together, the in vivo expression patterns of TTF-1, HNF-3, and Pax-8 proteins are sufficient to describe distinct anteroposterior regions of the mammalian

foregut axis (Fig. 7), and it is appealing to consider that these patterns may reflect regional codes for organ-specific lineage commitment along this axis, since the known target genes of these proteins are differentiated cell products.

A transcriptional hierarchy for HNF-3 and TTF-1. The expression studies discussed above suggest that in the lung bud, HNF-3 has a much earlier developmental onset than TTF-1. While TTF-1 is restricted to the anterior migrating edge of the lung diverticulum and is not detected in the floor of the pharynx (42), HNF-3 transcripts are seen morphologically earlier in the wall of the pharynx and, more intensely, in the laryngotracheal groove (51). Three additional lines of evidence support the idea that HNF-3 proteins occupy a primary position in the transcriptional hierarchy that leads to lineage commitment and cell-type-specific gene expression for endodermal derivatives. First, dedifferentiated hepatocyte cell lines and somatic cell hybrids typically lack expression of most known transcription factors for liver-specific genes (1, 39, 54). Peculiarly, HNF-3 proteins often continue to be expressed in the dedifferentiated cells (1, 39, 54), and in response to a differentiation signal, these cells express a cascade of other known liver transcription factors and progress to a differentiated hepatic cell type (1). Second, HNF-3β is capable of binding to and activating its own promoter (57) and could thus sustain its own expression as well as downstream transcription factor genes such as the TTF-1 gene. Third, McPherson et al. have recently shown that HNF-3 is capable of specifically modifying nucleosomal positioning within the liver-specific albumin enhancer (47). In their model, they propose that HNF-3 could disrupt the nucleosome structure and thereby enhance the binding of other liver-enriched factors that appear later in development to nearby binding sites. This may be a prototypic function of HNF-3 proteins, since the crystal structure of HNF-3γ-bound DNA is similar to the structure of the DNA binding domain of histone H5 (13), a variant of histone H1 (60), and disruption of H1 is proposed to alter nucleosome structure (26).

For these reasons, it is appealing to consider that HNF-3 proteins may have similar functional roles in activation of lung-specific genes. In the lung, the first lung-specific gene product, SPC, is detected in the primordial lung bud once it has invaded the surrounding mesenchyme (75), and this expression appears to coincide with the expression of TTF-1 in cells that are restricted to the anterior migrating edge of the lung bud. These events fit nicely into a model wherein local expression of TTF-1 is required subsequent to HNF-3 expression to activate or maintain transcription of lung-specific genes. An interesting comparison is that the onset of expression of TTF-1 appears analogous to the onset of expression of liver transcription factor HNF-1, since this factor is expressed subsequent to HNF-3 in liver development and only after endodermal cells have begun to invade the surrounding mesenchyme (4). In particular, this coincides with a marked increase in transcription of the liver-specific albumin gene (4, 11). These observations suggest that factors expressed downstream of HNF-3 proteins have a more local role in activation or maintenance of organ-specific gene transcription along the foregut axis. It would be particularly informative to determine if ectopic expression of these factors along this axis can alter lineage commitment in vivo. In this regard, it is noteworthy that homeotic transformations occur by ectopic or null expression of the vertebrate *Hox* gene family (9) and that ectopic expression of a single MyoD protein is sufficient to initiate skeletal muscle-specific gene expression in hearts of transgenic mice (49). For endodermal derivatives, these experiments may involve ectopic expression in the lung of factors that are normally restricted in expression to the thyroid or liver.

In conclusion, these data directly show that TTF-1 and HNF-3 are involved in lung-specific gene transcription, and this supports the idea that common factors control tissue-specific gene transcription along the foregut axis. The challenge that remains is to determine how these factors select from among possible target genes along this axis. Hypotheses that are based on combinatorial and context-dependent factor interaction as well as modification of these factors by posttranscriptional mechanisms suggest attractive future experiments designed to explicate this problem.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant HL-38859 and the Cystic Fibrosis Foundation (J.A.W.). R.J.B. was supported by NIH training grant HL-07527. R.D.L. was supported in part by a grant from the Associazione Italiana per la Ricera sul Cancro.

REFERENCES

1. **Ang, S.-L., A. Wierda, D. Wong, K. A. Stevens, S. Cascio, J. Rossant, and K. S. Zaret.** 1993. The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of the HNF3/*forkhead* proteins. Development **119:**1301–1315.
2. **Avvedimento, V. E., A. M. Musti, M. Ueffing, S. Obici, A. Gallo, M. Sanchez, D. DeBrasi, and M. E. Gottesman.** 1991. Reversible inhibition of a thyroid-specific *trans*-acting factor by Ras. Genes Dev. **5:**22–28.
3. **Bingle, C. D., and J. D. Gitlin.** 1993. Identification of hepatocyte nuclear factor-3 binding sites in the Clara cell secretory protein gene. Biochem. J. **295:**227–232.
4. **Blumenfeld, M., M. Maury, T. Chouard, M. Yaniv, and H. Condamine.** 1991. Hepatic nuclear factor 1 (HNF1) shows a wider distribution than products of its known target genes in the developing mouse. Development **113:**589–599.
5. **Bohinski, R. J., J. A. Huffman, J. A. Whitsett, and D. L. Lattier.** 1993. Cis-active elements controlling lung cell-specific expression of human pulmonary surfactant protein B gene. J. Biol. Chem. **268:**11160–11166.
6. **Bohinski, R. J., B. R. Stripp, T. R. Korfhagen, S. W. Glasser, and J. A. Whitsett.** Unpublished data.
7. **Borgmeyer, U., J. Nowock, and A. E. Sippel.** 1984. The TGGCA-binding protein: a eukaryotic nuclear protein recognizing a symmetrical sequence on double-stranded DNA. Nucleic Acids Res. **12:**4295–4311.
8. **Boshart, M., M. Klüppel, A. Schmidt, G. Schütz, and B. Luckow.** 1992. Reporter constructs with low background activity utilizing the *cat* gene. Gene **110:**129–130.
9. **Botas, J.** 1993. Control of morphogenesis and differentiation by HOM/Hox genes. Curr. Opin. Cell Biol. **5:**1015–1022.
10. **Cao, Z., R. M. Umek, and S. L. McKnight.** 1991. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3 cells. Genes Dev. **5:**1538–1552.
11. **Cascio, S., and K. S. Zaret.** 1991. Hepatocyte differentiation initiates during endodermal-mesenchymal interactions prior to liver formation. Development **113:**217–225.
12. **Civitareale, D., R. Lonigro, A. J. Sinclair, and R. Di Lauro.** 1989. A thyroid-specific nuclear protein essential for tissue-specific expression of the thyroglobulin promoter. EMBO J. **8:**2537–2542.
13. **Clark, K. L., E. D. Halay, E. Lai, and S. K. Burley.** 1993. Co-crystal structure of the HNF3/*forkhead* DNA recognition motif resembles histone H5. Nature (London) **364:**412–420.
14. **Clerc, R. G., L. M. Corcoran, J. H. LeBowitz, D. Baltimore, and P. A. Sharp.** 1988. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. Genes Dev. **2:**1570–1581.
15. **Clevidence, D. E., D. G. Overdier, W. Tao, X. Qian, L. Pani, E. Lai, and R. H. Costa.** 1993. Identification of nine tissue-specific transcription factors of the hepatocyte nuclear factor 3/forkhead DNA-binding-domain family. Proc. Natl. Acad. Sci. USA **90:**3948–3952.
16. **Costa, R. H., and D. R. Grayson.** 1991. Site-directed mutagenesis

of hepatocyte nuclear factor (HNF) binding sites in the mouse transthyretin (TTR) promoter reveal synergistic interactions with its enhancer region. Nucleic Acids Res. **19:**4139–4145.

17. **Costa, R. H., D. R. Grayson, and J. E. Darnell, Jr.** 1989. Multiple hepatocyte-enriched nuclear factors in the regulation of transthyretin and α1-antitrypsin genes. Mol. Cell. Biol. **9:**1415–1425.
18. **Damante, G., and R. Di Lauro.** 1991. Several regions of Antennapedia and thyroid transcription factor 1 homeodomains contribute to DNA binding specificity. Proc. Natl. Acad. Sci. USA **88:**5388–5392.
19. **Damante, G., D. Fabbro, L. Pellizzari, D. Civitareale, S. Guazzi, M. Polycarpou-Schwartz, S. Cauci, F. Quadrifoglio, S. Formisano, and R. Di Lauro.** Unpublished data.
20. **Damante, G., G. Tell, S. Formisano, D. Fabbro, L. Pellizzari, and R. Di Lauro.** 1993. Effect of salt concentration on TTF-1 HD binding to specific and non-specific DNA sequences. Biochem. Biophys. Res. Commun. **197:**632–638.
21. **Darnell, J. E.** 1982. Variety in the level of gene control in eukaryotic cells. Nature (London) **297:**365–371.
22. **Davis, R. L., H. Weintraub, and A. B. Lassar.** 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell **51:**987–1000.
23. **De Simone, V., and R. Cortese.** 1991. Transcriptional regulation of liver-specific gene expression. Curr. Opin. Cell Biol. **3:**960–965.
24. **Drewes, T., L. Klein-Hitpass, and G. U. Ryffel.** 1991. Liver specific transcription factors of the HNF3-, C/EBP-, and LFB1-families interact with the A-activator binding site. Nucleic Acids Res. **19:**6383–6389.
25. **Faisst, S., and S. Meyer.** 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res. **20:**3–26.
26. **Felsenfeld, G.** 1992. Chromatin as an essential part of the transcriptional mechanism. Nature (London) **355:**219–224.
27. **Francis-Lang, H., M. Price, M. Polycarpou-Schwartz, and R. Di Lauro.** 1992. Cell-type-specific expression of the rat thyroperoxidase promoter indicates common mechanisms for thyroid-specific gene expression. Mol. Cell. Biol. **12:**576–588.
28. **Francis-Lang, H., M. Zannini, M. De Felice, M. T. Berlingieri, A. Fusco, and R. Di Lauro.** 1992. Multiple mechanisms of interference between transformation and differentiation in thyroid cells. Mol. Cell. Biol. **12:**5793–5800.
29. **Gazdar, A. F., R. I. Linnoila, Y. Kurita, H. K. Oie, J. L. Mulshine, J. C. Clark, and J. A. Whitsett.** 1990. Peripheral airway cell differentiation in human lung cancer cell lines. Cancer Res. **50:**5481–5487.
30. **Grange, T., J. Roux, G. Rigaud, and R. Pictet.** 1990. Cell-type specific activity of two glucocorticoid responsive units of rat tyrosine aminotransferase gene is associated with multiple binding sites for C/EBP and a novel liver-specific nuclear factor. Nucleic Acids Res. **19:**131–139.
31. **Guazzi, S., M. Price, M. De Felice, G. Damante, M.-G. Mattei, and R. Di Lauro.** 1990. Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. EMBO J. **9:**3631–3639.
32. **Haugen, B. R., W. M. Wood, D. F. Gordon, and E. C. Ridgway.** 1993. A Thyrotrope-specific variant of Pit-1 transactivates the thyrotropin β promoter. J. Biol. Chem. **268:**20818–20824.
33. **Hennighausen, L., and H. Lubon.** 1987. Interaction of protein with DNA *in vitro*. Methods Enzymol. **152:**721–735.
34. **Ingraham, H. A., R. Chen, H. J. Mangalam, H. P. Elsholtz, S. E. Flynn, C. R. Lin, D. M. Simmons, L. Swanson, and M. G. Rosenfeld.** 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell **55:**519–529.
35. **Jackson, D. A., K. E. Rowander, K. Stevens, C. Jiang, P. Milos, and K. S. Zaret.** 1993. Modulation of liver-specific transcription by interactions between hepatocyte nuclear factor 3 and nuclear factor 1 binding DNA in close apposition. Mol. Cell. Biol. **13:**2401–2410.
36. **Jaenisch, R.** 1988. Transgenic animals. Science **240:**1468–1474.
37. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58:**799–839.
38. **Kapiloff, M. S., Y. Farkash, M. Wegner, and M. G. Rosenfeld.** 1991. Variable effects of phosphorylation of Pit-1 dictated by the DNA response elements. Science **253:**786–789.
39. **Kuo, C. J., P. B. Conley, L. Chen, F. M. Sladek, J. E. Darnell, Jr., and G. R. Crabtree.** 1992. A transcriptional hierarchy involved in mammalian cell-type specification. Nature (London) **355:**457–461.
40. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82:**488–492.
41. **Lai, E., V. R. Prezioso, W. Tao, W. S. Chen, 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.
42. **Lazzaro, D., M. Price, M. De Felice, and R. Di Lauro.** 1991. The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. Development **113:**1093–1104.
43. **Luo, Y., H. Fujii, T. Gerster, and R. G. Roeder.** 1992. A novel B cell-derived coactivator potentiates the activation of immunoglobin promoters by octamer-binding transcription factors. Cell **71:**231–241.
44. **MacGregor, G. R., and C. T. Caskey.** 1989. Construction of plasmids that express *E. coli* β-galactosidase in mammalian cells. Nucleic Acids Res. **6:**2365.
45. **MacGregor, G. R., G. P. Nolan, S. Fiering, M. Roederer, and L. A. Herzenberg.** 1991. Use of *E. coli lacZ* (β-galactosidase) as a reporter gene, p. 217–235. *In* E. J. Murray (ed.), Methods in molecular biology, vol. 7. Humana Press, Clifton, N.J.
46. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236:**1237–1244.
47. **McPherson, C. E., E.-Y. Shim, D. S. Friedman, and K. S. Zaret.** 1993. An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. Cell **75:**387–398.
48. **Milos, P. M., and K. S. Zaret.** 1992. A ubiquitous factor is required for C/EBP-related proteins to form stable transcription complexes on an albumin promoter segment in vitro. Genes Dev. **6:**991–1004.
49. **Miner, J. H., J. B. Miller, and B. J. Wold.** 1992. Skeletal muscle phenotypes initiated by ectopic MyoD in transgenic mouse heart. Development **114:**853–860.
50. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245:**371–378.
51. **Monaghan, P. A., K. H. Kaestner, E. Grau, and G. Schutz.** 1993. Postimplantation expression patterns indicate a role for the mouse *forkhead*/HNF-3α,β and γ genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. Development **119:**567–578.
52. **Nerlov, C., and E. B. Ziff.** 1994. Three levels of functional interaction determine the activity of the CCAAT/enhancer binding protein-α on the serum albumin promoter. Genes Dev. **8:**350–362.
53. **Nitsch, D., M. Boshart, and G. Schütz.** 1993. Activation of the tyrosine aminotransferase gene is dependent on synergy between liver-specific and hormone responsive elements. Proc. Natl. Acad. Sci. USA **90:**5479–5483.
54. **Nitsch, D., M. Boshart, and G. Schütz.** 1993. Extinction of tyrosine aminotransferase gene activity in somatic cell hybrids involves modification and loss of several essential transcriptional activators. Genes Dev. **7:**308–319.
55. **O'Reilly, M. A., A. F. Gazdar, R. E. Morris, and J. A. Whitsett.** 1988. Differential effects of glucocorticoid on expression of surfactant proteins in a human lung adenocarcinoma cell line. Biochim. Biophys. Acta **970:**194–204.
56. **Overdier, D. G., A. Porcella, and R. H. Costa.** 1994. The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino acid residues adjacent to the recognition helix. Mol. Cell. Biol. **14:**2755–2766.
57. **Pani, L., X. Qian, D. Clevidence, and R. H. Costa.** 1993. The restricted promoter activity of the liver transcription factor hepatocyte nuclear factor 3β involves a cell specific factor and positive autoactivation. Mol. Cell. Biol. **12:**552–562.
58. **Phelps, D. S., and J. Floros.** 1991. Localization of pulmonary surfactant proteins using immunohistochemistry and tissue in situ

hybridization. Exp. Lung Res. **17:**985–995.
59. **Plachov, D. K., K. Chowdhury, C. Walther, D. Simon, J. L. Guenet, and P. Gruss.** 1990. Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland. Development **110:**643–651.
60. **Ramakrishnan, V., J. T. Finch, V. Granziano, P. L. Lee, and R. M. Sweet.** 1993. Globular domain of histone H5 and its implications for nucleosome binding. Nature (London) **362:**219–223.
61. **Rigaud, G., J. Roux, R. Pictet, and T. Grange.** 1991. In vivo footprinting of the rat TAT gene: dynamic interplay between the glucocorticoid receptor and a liver-specific factor. Cell **67:**977–986.
62. **Rosenthal, N.** 1987. Identification of regulatory elements of cloned genes with functional assays. Methods Enzymol. **152:**704–720.
63. **Rupp, R. A. W., U. Kruse, G. Multhaup, U. Gobel, K. Beyreuther, and A. E. Sippel.** 1990. Chicken NF1/TGCCA proteins are encoded by at least three independent genes: NF1/A, NF1-B, and NF1-C with homologies in mammalian genomes. Nucleic Acids Res. **18:**2607–2616.
64. **Sasaki, H., and B. L. M. Hogan.** 1993. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. Development **118:**47–59.
65. **Sawaya, P. L., B. R. Stripp, J. A. Whitsett, and D. S. Luse.** 1993. The lung-specific CC10 gene is regulated by transcription factors from the AP-1, octamer, and hepatocyte nuclear factor 3 families. Mol. Cell. Biol. **13:**3860–3871.
66. **Schreiber, E., P. Matthias, M. M. Muller, and W. Schaffner.** 1989. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucleic Acids Res. **17:**6419.
67. **Shaw-White, J., and J. A. Whitsett.** Unpublished data.
68. **Simmons, D. M., J. W. Voss, H. A. Ingraham, J. M. Holloway, R. S. Broide, M. G. Rosenfeld, and L. W. Swanson.** 1990. Pituitary cell phenotypes involve cell-specific Pit-1 mRNA translation and synergistic interactions with other classes of transcription factors. Genes Dev. **4:**695–711.
69. **Stahlman, M. T., and J. A. Whitsett.** Unpublished data.
70. **Stripp, B. R., J. A. Huffman, and R. J. Bohinski.** Structure and regulation of the murine Clara cell secretory protein gene. Genomics **20:**27–35.
71. **Tsai, S.-F., D. I. K. Martin, L. I. Zon, A. D. D'Andrea, G. G. Wong, and S. H. Orkin.** 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. Nature (London) **339:**446–451.
72. **Weaver, T. E., and J. A. Whitsett.** 1991. Function and regulation of expression of pulmonary surfactant-associated proteins. Biochem. J. **273:**249–264.
73. **Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lassar.** 1991. The *myo*D gene family: nodal point during specification of the muscle cell lineage. Science **251:**761–766.
74. **Weintraub, H., S. J. Tapscott, R. L. Davis, M. J. Thayer, M. A. Adam, A. B. Lassar, and D. A. Miller.** 1989. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. Proc. Natl. Acad. Sci. USA **86:**5434–5438.
75. **Wert, S. E., S. W. Glasser, T. R. Korfhagen, and J. A. Whitsett.** 1993. Transcriptional elements from the human SP-C gene direct expression in the primordial respiratory epithelium of transgenic mice. Dev. Biol. **156:**426–443.
76. **Wikenheiser, K. A., D. K. Vorbroker, W. R. Rice, J. C. Clark, C. J. Bachurski, H. K. Oie, and J. A. Whitsett.** 1993. Production of immortalized distal respiratory epithelial cell lines from surfactant protein C/simian virus 40 large tumor antigen transgenic mice. Proc. Natl. Acad. Sci. USA **90:**11029–11033.
77. **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.
78. **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.
79. **Zannini, M., H. Francis-Lang, D. Plachov, and R. Di Lauro.** 1992. Pax-8, a paired domain-containing protein, binds to a sequence overlapping the recognition site of a homeodomain and activates transcription from two thyroid-specific promoters. Mol. Cell. Biol. **12:**4230–4241.