Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 3948–3952, May 1993
Biochemistry

Identification of nine tissue-specific transcription factors of the hepatocyte nuclear factor 3/forkhead DNA-binding-domain family

(tissue-specific transcription factors/gene family/differentiation)

DEREK E. CLEVIDENCE*, DAVID G. OVERDIER*, WUFAN TAO†, XIAOBING QIAN*, LUCA PANI*‡, ESENG LAI†, AND ROBERT H. COSTA*§

*Department of Biochemistry (M/C 536), University of Illinois College of Medicine, Chicago, IL 60612; and †Division of Endocrinology and Program of Cell Biology and Genetics, Memorial Sloan–Kettering Cancer Center, New York, NY 10021

Communicated by Irving M. Klotz, January 4, 1993

ABSTRACT **Hepatocyte nuclear factor (HNF)-3α, -3β, and -3γ are liver transcription factors that mediate the coordinate expression of a number of hepatocyte-specific genes. The HNF-3 proteins share DNA-binding-domain homology among themselves and with the *Drosophila* homeotic protein forkhead (fkh). The HNF-3/fkh DNA-binding domain constitutes an uncharacterized protein motif that recognizes its cognate DNA binding site as a monomer. Additional HNF-3/fkh-related proteins are known to be required for determination events during embryogenesis in *Drosophila* and *Xenopus*. In this report, we describe the isolation of nine additional HNF-3/fkh homologue (HFH) clones from rodent tissue cDNAs by using both low-stringency hybridization and a polymerase chain reaction protocol. Many of the HFH genes exhibit a tissue-restricted expression pattern and are transcribed in tissues other than liver, including brain, kidney, lung, and intestine. The HNF-3/fkh motif therefore comprises a large gene family of transcription factors that play a role in tissue-specific gene regulation and development.**

Cellular differentiation, histogenesis, and development are the consequences of differential gene expression resulting from temporal cascades of gene activation or repression (1, 2). In eukaryotes such gene activation is regulated principally at the transcriptional level by the coordinated activity of trans-acting factors that recognize cis-acting DNA sequences in promoter and enhancer regions (3, 4). Cloning and subsequent characterization have revealed that these transcriptional regulatory factors are modular in structure and consist of independently functioning domains (5–8). The transcription factor domains include those involved in specific DNA recognition (DNA-binding domains), the formation of homodimeric or heterodimeric proteins (dimerization domains), and the stimulation of RNA polymerase II initiation (activation domains). Each transcription factor draws specificity from its DNA-binding domain, which allows the factor to recognize only promoters containing the cognate DNA binding sites. The isolation of related family members by low-stringency hybridization with DNA-binding-domain probes has demonstrated that transcription factors have evolved as gene families in which the DNA-binding domain, the transcriptional activation domain, or a dimerization domain has been conserved. Development of methods to determine high-affinity DNA binding sites of transcription factor family members has also allowed the identification of their putative target promoters (9, 10). The availability of related transcription factor probes has facilitated the analysis of family members' expression patterns. This, in turn, has provided a broader understanding of their roles in transcriptional regulation during differentiation, development, and disease. A general grouping of the known transcription factor families would include those with the helix–turn–helix motif (e.g., homeodomain and POU-homeodomain proteins), the zinc finger motif (e.g., Sp1, and the steroid, thyroid hormone, and retinoic acid receptor superfamily), the basic DNA-binding domain and helix–loop–helix and leucine zipper dimerization motifs (e.g., C/EBP, c-Jun, c-Fos, MyoD, and Myc family), the ankyrin dimerization motif (e.g., NF-κB, GA binding protein, and Ets family), and the paired-box motif (e.g., murine Pax family) (11–23).

Recent evidence suggests that another class of transcription factors, sharing a DNA-binding domain of the type found in hepatocyte nuclear factor (HNF)-3α, 3β, and 3γ and in the *Drosophila* protein forkhead (fkh), functions critically in cellular differentiation and development (24–27). The rat liver transcription factors HNF-3α, -3β, and -3γ cooperate with several other liver-enriched factors to regulate the restricted expression of several genes important for liver function (28–31). HNF-3 is also involved in regulating the expression of another liver regulatory protein, HNF-1, and is therefore known to function at an early position within the hierarchy of factors involved in hepatocyte differentiation (32). Further evidence for involvement of HNF-3-related proteins in development is provided by the homeotic protein, fkh, which is critical for cellular determination of intestinal structures in *Drosophila melanogaster* (26). In addition to the DNA-binding domain, fkh also shares homology within the HNF-3 activation domain, suggesting that HNF-3 may be its mammalian homologue (25, 27, 33). Several other HNF-3/fkh family members participating in determination events during embryogenesis have recently been identified, including those encoded by the sloppy paired loci (slp1 and slp2) and an activin-inducible *Xenopus laevis* gene, XFKH1 (34, 35). The isolation of three HNF-3/fkh clones in nonhepatic adult tissues (BF-1, ILF, and HTLF; refs. 36–38) suggests that the HNF-3/fkh motif may comprise a larger regulatory family involved in cell-type-specific gene regulation (Fig. 1). Finally, the HNF-3/fkh motif represents an evolutionally ancient domain among eukaryotes because the HNF-3/fkh is conserved in the yeast *Saccharomyces cerevisiae* (39).

In this study, we initially isolated four HNF-3/fkh-related genes by low-stringency hybridization screening of rodent brain cDNA, lung cDNA, and genomic libraries with the HNF-3/fkh motif probe. We developed a polymerase chain reaction (PCR) protocol using degenerate primers synthesized to conserved sequences with the DNA-binding domain

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HNF, hepatocyte nuclear factor; fkh, forkhead; HFH, HNF-3/fkh homologue.

‡Present address: The Fidia–Georgetown Institute of Neurosciences, Georgetown University Medical School, Washington, DC 20007.
§To whom reprint requests should be addressed.

that allowed the isolation of five additional HNF-3/fkh homologue (HFH) clones from tissue cDNAs.¶ This PCR protocol can be used to amplify HNF-3/fkh family members from cDNAs prepared from a variety of adult or embryonic tissues. Many of the HFH genes exhibit a tissue-restricted expression pattern and possess DNA-recognition properties which differ from those of the HNF-3 proteins. These results implicate the HNF-3/fkh family as critical regulatory proteins for tissue-specific gene regulation and differentiation.

MATERIALS AND METHODS

Low-Stringency Library Screening and PCR Amplification. HFH-1 was isolated from a rat lung cDNA library (Clontech) by low-stringency hybridization with an HNF-3β DNA probe encoding the DNA-binding domain. Hybridization of nitrocellulose lifts was done in 6× standard saline citrate (SSC)/5× Denhardt's solution/50 mM sodium phosphate, pH 6.8/0.05% sodium pyrophosphate/0.2% SDS containing denatured sonicated salmon sperm DNA (50 μg/ml) at 50°C for 36 hr. The filters were washed at 50°C in successive washes of 6× SSC and 2× SCC containing 0.05% sodium pyrophosphate and 0.2% SDS. HFH-B1 and HFH-B2 were obtained from a mouse brain cDNA library by low-stringency hybridization to the HNF-3 DNA-binding-domain probe according to the protocol above. Recombinant phage DNA was prepared from an agarose plate lysate (40).

We have used comparisons among the HNF-3, HFH-1, HFH-B2, HFH-B3, and BF-1 motifs to synthesize degenerate primers corresponding to conserved residues within the DNA-binding domain. We synthesized a sense primer corresponding to the KPPYSYI amino acid sequence [5′-CGC-AAGCTT-AAR-CC(T/C/A)-CC(T/C/A)-TA(T/A)-TCN-TAY-AT-3′] and an antisense primer to WQNSIRH [5′-GCGGTCGAC-RTG-YC(G/T)-RAT-NGA-RTT-CTG-CCA-3′] that contained *Hin*dIII and *Sal* I restriction sites at the 5′ end (underlined) and used them for PCR amplification of HNF-3/fkh family members (indicated in Fig. 1). Total RNA was prepared from dissected rat tissues by guanidinium thiocyanate/phenol/chloroform extraction (41) and polyadenylylated RNA was selected on oligo(dT)-cellulose. Poly(A)$^+$ mRNA (5 μg) prepared from various rat tissues was primed with oligo(dT) for cDNA synthesis by reverse transcriptase (Superscript; Bethesda Research Laboratories). The cDNA was used as a template for PCR amplification with the degenerate primers and Vent DNA polymerase (New England Biolabs) for 34 cycles at 48°C hybridization temperature. To facilitate restriction enzyme digestion, the PCR products (180 nucleotides long) were differentially precipitated from the primers in ice-cold 10% (wt/vol) polyethylene glycol/0.8 M NaCl. The PCR products were digested with *Hin*dIII and *Sal* I, gel-purified, and ligated into the corresponding sites of pGEM-1 (Promega), and then the transformed bacterial colonies were screened by high-stringency hybridization with the HNF-3α and HNF-3β DNA-binding-domain probes (40). Weakly hybridizing clones that represented novel PCR products were selected for dideoxy DNA sequencing analysis (42) using the Sequenase enzyme (United States Biochemical). Unique PCR probes were used to screen genomic and cDNA libraries at high stringency and isolate HNF-3/fkh family members (65°C hybridization; final washes in 0.2× SSC). The DNA-binding-domain sequence was determined with specific oligonucleotide primers.

Site-Directed Mutagenesis and Gel Shift Assays. Mutations within the HNF-3γ DNA-binding domain were introduced by oligonucleotide-directed mutagenesis (43). RNA was synthesized from the HNF-3γ mutant template by T7 phage RNA polymerase and used to program the synthesis of mutant protein in a rabbit reticulocyte lysate system (Promega). Incorporation of [^{35}S]methionine was determined by the method described by Promega and was used to normalize the amount of HNF-3γ protein produced from the *in vitro* translation reactions. Equal amounts of each of the mutant HNF-3γ proteins were incubated with labeled double-stranded oligonucleotide containing the high-affinity HNF-3 recognition site (transthyretin promoter, positions −111 to −85; ref. 28) and protein–DNA complex formation was visualized by gel shift assays (25). Specific protein–DNA complexes were determined by competition with a 100-fold excess of unlabeled HNF-3 binding-site oligonucleotide.

Northern Blots and T2 RNase Protection Assays. Northern blots were made by fractionating 10 μg of total RNA in denaturing formaldehyde agarose gels and then transferring the RNA to nitrocellulose (25). Antisense RNA probes were prepared from *Hin*dIII-linearized HFH PCR plasmids with T7 RNA polymerase and then used in T2 RNase protection assays with total RNA (28).

RESULTS

Isolation of the HFH DNA-Binding Domains. Low-stringency library screening and PCR amplification of cDNA templates prepared from a variety of tissue RNAs led to the cloning of an additional nine members of the HFH family. Low-stringency screening of cDNA libraries with the HNF-3/fkh DNA-binding motif provided two additional brain clones (HFH-B2 and HFH-B3) and one lung cDNA clone (HFH-1) (Fig. 1). Comparison of the DNA-binding domains of these HNF-3/fkh family members allowed us to design two degenerate primers corresponding to the conserved KPPYSYI and WQNSIRH sequences (primer; Fig. 1) that were used to generate PCR products from heart, lung, brain, kidney, and intestine (Caco-2 human colon carcinoma cells) cDNAs. After cloning and sequencing the HFH PCR products, we further characterized those which exhibited the greatest amino acid diversity from the original HNF-3/fkh motif (Fig. 1). To isolate the remaining HFH DNA-binding domains, we screened rat genomic and cDNA libraries at high stringency with each of the HFH PCR products. We completed the HFH motifs' DNA sequence with specific oligonucleotides and used this to derive the encoded amino acid sequence. The alignment of the DNA-binding domains of the HFH proteins with other published HNF-3/fkh family members is shown in Fig. 1. The comparison is made with respect to the HNF-3α DNA-binding domain and is presented in order of decreasing homology with HNF-3α sequence.

Comparison of the HFH DNA-binding domains allows us to identify conserved sequences whose retention may be required for the basic structure of the HFH motif. The alignment also identifies three punctuated divergent regions at the amino terminus (Fig. 1, regions 1–3). More significant sequence variation is observed at the carboxyl terminus, especially within a stretch of basic amino acid residues (Fig. 1, region 4). The divergent HNF-3/fkh family members BF-1 (36), ILF and HTLF (37, 38), and HFH-1 (data not shown) possess DNA-recognition specificities that differ significantly from those of HNF-3. The HNF-3 DNA-recognition sequence (12), C(A/T)AARTCAATA, will not compete for ILF or HTLF binding to their purine-rich DNA sites, GGAGRARRR (38). Therefore, the sequence divergence that is apparent within the HFH DNA-binding domains may mediate this recognition of different DNA sites. The altered DNA recognition sites of the HFH proteins will allow them to regulate the transcription of target promoters that differ from those of the liver HNF-3 proteins.

¶The nucleotide sequences of the HFH DNA-binding domains have been deposited in the GenBank data base (accession nos. L13201–L13207, L13192, and L13193).

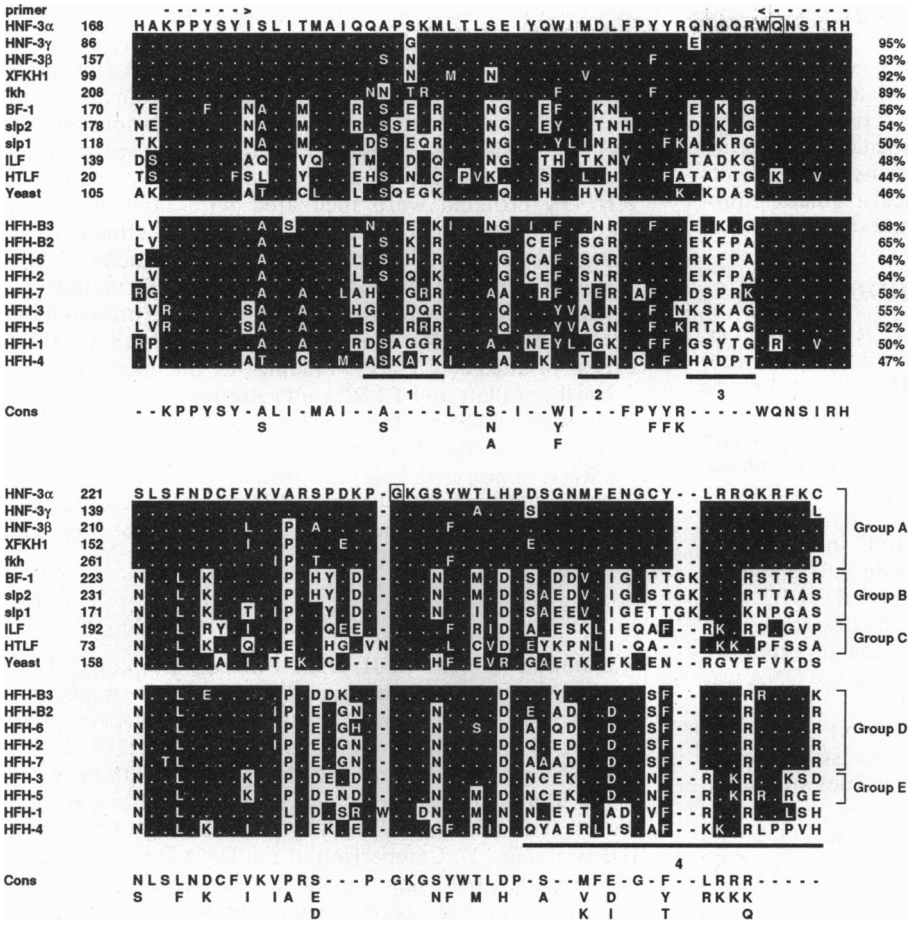

FIG. 1. Sequence alignment of the HFH DNA-binding domains. The amino acid sequence (one-letter code) of the HNF-3α (24) DNA binding domain is used as a basis for comparison with the HFH sequences, which are presented in decreasing order of homology (divergent regions are highlighted by bars 1–4). Conserved amino acid residues are highlighted by the shaded region (white dots designate identities; white letters indicate conservative changes), gaps have been introduced to maximize homology (dashes), boxed residues indicate the position of the HFH-3, -4, -5 introns, and a loose consensus sequence (Cons) of the HNF-3/fkh family is shown. Also indicated are the percent identity with respect to the HNF-3α DNA-binding domain, and the related HFH groups A–E are indicated by brackets. The published HNF-3 family members and their amino acid positions (indicated by numbers) are derived from the following: HNF-3β and -3γ (25), fkh (26), XFKH1 (*Xenopus* forkhead) (34), slp1 and slp2 (sloppy paired locus, *Drosophila*) (35), ILF (interleukin binding factor) and HTLF (human T-cell leukemia virus enhancer factor) (37, 38), BF-1 (brain factor 1) (36), and the fkh homologue in yeast (39). The HFH clones were isolated by using degenerate primers corresponding to the conserved KPPYSYI and WQNSIRH regions for PCR with cDNA templates from various rat tissues.

In addition to the strong identity between the HNF-3 and fkh DNA-binding domains, the HFH proteins can be arranged in HNF-3/fkh-related groups (Fig. 1, groups A–E). The *Drosophila* slp1 and slp2 proteins show considerable similarity to BF-1 yet diverge within the basic residues at the carboxyl terminus (Fig. 1, group B). Group C is defined by the ILF and HTLF factors, and group D comprises five new HFH clones, three of which exhibit significant conservation within divergent regions 1–4 (Fig. 1, HFH-B2, HFH-6, and HFH-2). Furthermore, analysis of the HFH-3 and HFH-5 genomic structure revealed that their DNA-binding domains are interrupted at the same position by intronic sequences (boxed G residue) and that they define the final related group, E (Fig. 1). The HFH-1 and HFH-4 motifs are two family members which do not fall within any related group, and perhaps additional clones exist that share homology with these two HFH domains. Moreover, the HFH DNA-binding domains that were encoded by two different exons, including ILF, HFH-3, HFH-4 (boxed Q residue), and HFH-5, exhibited only limited conservation with the HNF-3α sequence. On the other hand, the DNA-binding domains of the HFH group D were contained within a single exon, as were the DNA-binding domains encoded by the HNF-3 genes (44), and these HFH members exhibited greater homology to the HNF-3 sequence. These results suggest that exon shuffling may play a role in generating diversity at the carboxyl terminus of the HFH DNA-binding motif. Finally, the HFH-1 motif contains an insertion within the conserved PGKG residues and shows variation in the WQNSIR sequence (Fig. 1). In contrast to the aforementioned members, the diversity within the HFH-1 motif is not the result of splicing, and therefore HFH-1 may define a subgroup within the HNF-3/fkh family.

Analysis of Mutants Within the HNF-3γ DNA-Binding Domain. We prepared a series of amino acid substitutions within the HNF-3γ DNA-binding domain in order to identify the regions which are necessary for recognition of the HNF-3 DNA site. The HNF-3γ mutant proteins were expressed by translation in a rabbit reticulocyte lysate system and then assessed for protein–DNA complex formation by gel shift assay using the HNF-3 DNA binding site (transthyretin promoter). We compared equal amounts of wild-type and mutant HNF-3γ protein for DNA-binding activity by normalizing each translation reaction to the [^{35}S]methionine radioactivity incorporated into protein (see *Materials and Methods*). In addition, we included a competition lane in our gel shift assays to facilitate the identification of the HNF-3γ-specific protein–DNA complex. The binding activities exhibited by the HNF-3γ mutant proteins and their corresponding sequences are summarized in Fig. 2. Amino acid substitutions within conserved regions of the HNF-3γ DNA-binding domain resulted in proteins that were poorly bound to the HNF-3 DNA site, thus confirming a functional role for DNA recognition (Fig. 2; mutants M1–3 and M6–8). In contrast, we observed normal HNF-3 binding activity when substitutions were made within either region 2 or the amino portion of region 4 (Fig. 2; M4 and M9). This suggests that these divergent amino acid residues may not be critical for recognition of the HNF-3 DNA site. However, sequences within either region 3 or the basic stretch in region 4 abrogated binding to the HNF-3 DNA site, suggesting involvement in specific DNA recognition (Fig. 2; M5 and M10). Therefore, because the sequences within regions 3 and 4 vary among the HFH DNA-binding proteins, it is tempting to speculate that these residues may bestow different DNA-binding specificities upon each of the HFH-related groups. Another interpretation of the data is that these divergent

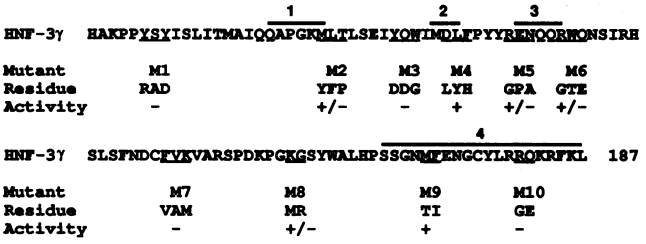

FIG. 2. Summary of DNA-binding activity exhibited by proteins with site-directed mutations within the HNF-3γ DNA-binding motif. Shown are the HNF-3γ DNA-binding domain sequence (ref. 25; position within protein is indicated by number), the regions that displayed sequence divergence among the HFH family members (indicated by bars 1–4 as in Fig. 1), the amino acid residues targeted by the mutation (underlined sequences), the residues that were substituted by the mutation (shown below M), and DNA-binding activity of the HNF-3γ mutant protein (+, wild-type activity; ±, 20% of wild-type activity; −, no detectable DNA binding). Specific mutations were introduced into the HNF-3γ DNA-binding-domain sequence by oligonucleotide-directed mutagenesis as described (28, 29, 33). The HNF-3γ mutant protein was synthesized in a rabbit reticulocyte *in vitro* translation system that was programmed by RNA synthesized *in vitro* by T7 RNA polymerase using mutated HNF-3γ template (25). Protein–DNA complex formation was analyzed by gel shift assay using the transthyretin promoter HNF-3 site (nucleotides −111 to −85) that possessed strong affinity for the HNF-3 proteins (28).

sequences confer structural stability and may not be involved in the actual DNA-binding specificity.

HFH mRNAs Display Tissue-Restricted Distribution. The tissue-restricted expression pattern of several HFH genes suggests that they may play an important role in the regulation of cell-type-specific genes. The HFH adult mRNA expression levels were examined by Northern blot analysis using RNA prepared from various rat tissues (Fig. 3*A*). The HFH-1, HFH-3, HFH-4, and HFH-5 mRNAs were restricted in their cellular distribution and in most cases were transcribed in tissues other than liver. In contrast, the HFH-2 gene showed expression in all tissues examined except for spleen, and HFH-7 mRNA, which was isolated by genomic library screening, was not detectable in the tissues we examined (Table 1). The HFH-3 gene displayed a kidney-specific expression pattern; the HFH-4 gene was abundantly expressed in the lung, with lower expression in the brain. Moderate expression was observed for HFH-1 in lung and kidney. Low-level mRNA expression was detected by T2 RNase protection for HFH-5 in lung and liver, for HFH-6 in brain, lung, and intestine, and for HFH-B2 and HFH-B3 in brain (Fig. 3*B* and data not shown). Interestingly, several of the HFH genes were transcribed in the lung, a tissue also expressing the HNF-3α and -3β genes (25).

DISCUSSION

The HNF-3/fkh motif constitutes a family of transcription factors that are critical participants in the regulation of cell-type-specific genes. Regions of sequence divergence within the HNF-3/fkh motif may allow each HFH group to activate different target promoters through interaction with distinct DNA elements. The utilization of related binding motifs that allow varied DNA site recognition in proteins regulating cellular specialization is evident in other transcription factor families as well (1–4, 11–23). For example, the identification and subsequent expression studies of homeodomain family members in *Drosophila* and mammals (POU-homeodomain and Hox genes) revealed their target genes and role in developmental regulatory cascades (1, 11–13, 23). The role of the steroid hormone, thyroid hormone, and retinoic

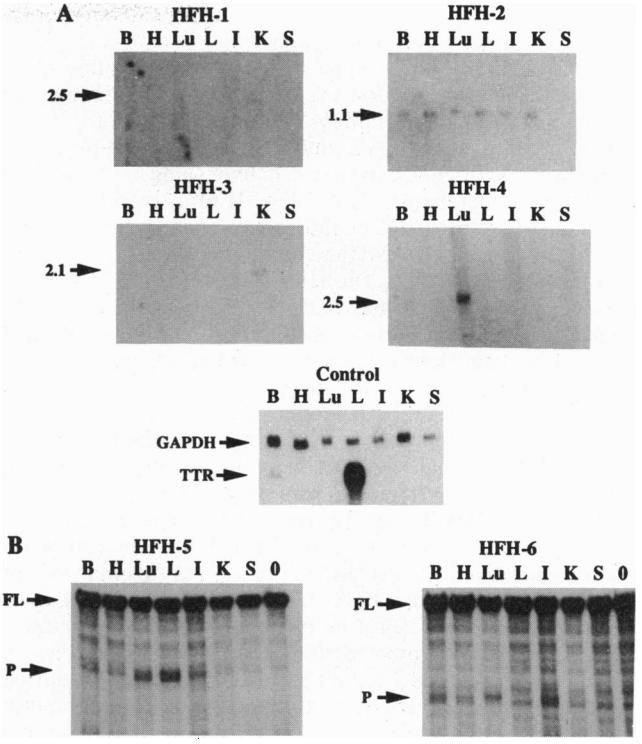

FIG. 3. Several HFH clones demonstrate a restricted expression pattern in tissues other than liver. (*A*) Northern blots with HFH probes. Samples (10 μg) of total RNA isolated from adult rat brain (B), heart (H), lung (Lu), liver (L), intestine (I; duodenum), kidney (K), and spleen (S) were fractionated in a denaturing (formaldehyde) agarose gel and transferred to nitrocellulose for Northern blot analysis with the indicated HFH probes (24- to 48-hr exposure). An arrow indicates the position of the hybridizing mRNA, and the mRNA size is given in kilobases. Intact tissue RNA was demonstrated by probing blots with the rat glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and transthyretin (TTR) cDNAs (45). (*B*) T2 RNase protection with the HFH-5 and HFH-6 probes. Total RNA from the indicated tissues (as in *A*) was hybridized with HFH-5 or HFH-6 antisense RNA probes, followed by digestion with T2 RNase and fractionation in a sequencing gel (28, 29). FL indicates the migration of the undigested RNA probe, and P shows the position of the RNase resistant product. Lane 0 is a control that included hybridization of HFH probes with yeast tRNA followed by T2 RNase analysis.

acid receptor family in embryogenesis has similarly been demonstrated through expression analysis and target-tissue identification (14–17). The availability of HNF-3/fkh family clones will enable us to perform similar expression studies to establish the significance of HFH proteins in cellular differentiation.

Although sequence variation confers DNA binding-site heterogeneity within transcription factor families, the regulatory function of subgroups that recognize identical sites can

Table 1. Summary of tissue expression pattern of HFH mRNA

	B	H	Lu	L	I	K	S
HFH-1	−	−	+	−	−	+	−
HFH-2	++	++	++	++	++	++	−
HFH-3	−	−	−	−	−	++	−
HFH-4	+	−	+++	−	−	−	−
HFH-5	−	−	+	+	−	−	−
HFH-6	+	±	+	−	+	−	−

HFH mRNA tissue distribution pattern was determined by Northern blot and T2 RNase analysis presented in Fig. 3. Tissues are abbreviated as in Fig. 3. RNA expression levels are indicated as follows: −, undetectable; ± or +, low; ++, moderate; +++, high.

be compounded through their differential expression during cellular determination and through their specific integration within cellular signaling systems. Implicitly, motifs other than the DNA-binding domain can functionally distinguish members of a regulatory family. Particularly illustrative is the C/EBP family, which is essential for adipocyte differentiation as well as for gene expression in liver, lung, and intestine (46–49). One member, C/EBPβ (NF-IL6), is responsible for interleukin 6 induction of hepatic acute-phase genes (49) and is also responsive to cAMP and calmodulin-mediated signals in other tissues (50, 51). The HNF-3 proteins will recognize identical binding sites and may therefore allow varied responses to different signal transduction pathways. Hence, subtle sequence variations can endow each family member with variable signal-coordinating activities. An interesting possibility is that HFH proteins could either activate or repress transcription consequential to such stimuli. Many of the HFH functional qualities may be resolved following characterization of the entire protein.

The HNF-3/fkh motif comprises one of the largest families of transcription factors involved in cellular determination events. Moreover, during the preparation of this manuscript five additional fkh-domain (FD) genes were isolated from *Drosophila* genomic libraries by low-stringency hybridization, and all differ from the HFH sequences defined herein (52). These FD genes were all expressed exclusively during embryonic development and exhibited a variety of temporal and positional expression patterns. Interestingly, the FD genes 3–5, which are expressed in cells of neuronal lineage, are most similar to the HFH group D proteins, several of which are also expressed in the brain. The *Drosophila* FD genes may also correspond to genes in mammals which serve as developmental regulators. The addition of nine HFH proteins by the work reported here has substantiated our belief that the HNF-3/fkh domain defines an extensive gene family, now consisting of 25 members, which participate in tissue-specific and developmental gene regulation.

Equal contribution was made by D.E.C. and D.G.O. We thank Susan Ross, Robert Storti, and Pradip Raychaudhuri for critically reviewing this manuscript; A. Tyner for the gift of Caco-2 intestinal cell RNA; and V. Sukhatme for the kidney cDNA library. This work was supported by grants from the National Institute of General Medical Sciences (GM43241) and the Council for Tobacco Research (2822) to R.H.C., and by an American Cancer Society grant (CD-494) to E.L.

1. McGinnis, W. & Krumlauf, R. (1992) *Cell* **68,** 283–302.
2. Levine, M. & Manley, J. L. (1989) *Cell* **59,** 405–408.
3. Johnson, P. & McKnight, S. L. (1989) *Annu. Rev. Biochem.* **58,** 799–839.
4. Mitchell, P. J. & Tjian, R. (1989) *Science* **245,** 371–378.
5. Hope, A. & Struhl, K. (1986) *Cell* **46,** 885–894.
6. Hollenberg, S. M. & Evans, R. (1988) *Cell* **55,** 899–906.
7. Ptashne, M. & Gann, A. F. (1990) *Nature (London)* **346,** 329–331.
8. Frankel, A. D. & Kim, P. S. (1991) *Cell* **65,** 717–719.
9. Blackwell, T. K. & Weintraub, H. (1990) *Science* **250,** 1104–1110.
10. Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N. & Weintraub, H. (1990) *Science* **250,** 1149–1151.
11. Kessel, M. & Gruss, P. (1990) *Science* **249,** 374–379.
12. Karin, M., Castrillo, J. & Theill, L. E. (1990) *Trends Genet.* **6,** 92–96.
13. Rosenfeld, M. G. (1991) *Genes Dev.* **5,** 897–907.
14. Evans, R. M. (1988) *Science* **240,** 889–893.
15. Zelent, A., Krust, A., Petkovich, M., Kastner, P. & Chambon, P. (1989) *Nature (London)* **339,** 714–717.
16. Dolle, P. E., Ruberte, E., Kastner, P., Petkovich, M., Stoner, C. M., Gudas, L. J. & Chambon, P. (1989) *Nature (London)* **342,** 702–705.
17. Mangelsdorf, D. J., Borgmeyer, U., Heyman, R. A., Zhou, J. Y., Ong, E. S., Oro, A. E., Kakizuka, A. & Evans, R. M. (1992) *Genes Dev.* **6,** 329–344.
18. Murre, C., McCaw, P. S. & Baltimore, D. (1989) *Cell* **56,** 777–783.
19. Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T. K., Turner, D., Rupp, R., Hollenberg, S., Zhuang, Y. & Lassar, A. (1991) *Science* **251,** 761–766.
20. Ghosh, S., Gifford, A. M., Riviere, L. R., Tempst, P., Nolan, G. P. & Baltimore, D. (1990) *Cell* **62,** 1019–1029.
21. Thompson, C. C., Brown, T. A. & McKnight, S. L. (1991) *Science* **253,** 762–768.
22. Macleod, K., Leprince, D. & Stehelin, D. (1992) *Trends Biochem. Sci.* **17,** 251–256.
23. Gruss, P. & Walther, C. (1992) *Cell* **69,** 719–722.
24. Lai, E., Prezioso, V. R., Smith, E., Litvin, O., Costa, R. H. & Darnell, J. E., Jr. (1990) *Genes Dev.* **4,** 1427–1436.
25. Lai, E., Prezioso, V. R., Tao, W., Chen, W. & Darnell, J. E., Jr. (1991) *Genes Dev.* **5,** 416–427.
26. Weigel, D., Jurgens, G., Kuttner, F., Seifert, E. & Jackle, H. (1989) *Cell* **57,** 645–658.
27. Weigel, D. & Jackle, H. (1990) *Cell* **63,** 455–456.
28. Costa, R. H., Grayson, D. R. & Darnell, J. E., Jr. (1989) *Mol. Cell. Biol.* **9,** 1415–1425.
29. Costa, R. H. & Grayson, D. R. (1991) *Nucleic Acids Res.* **19,** 4139–4145.
30. Lai, E. & Darnell, J. E., Jr. (1991) *Trends Biochem. Sci.* **16,** 427–430.
31. Liu, J., DiPersio, C. M. & Zaret, K. S. (1991) *Mol. Cell. Biol.* **11,** 773–784.
32. Kuo, C. J., Conley, P. B., Chen, L., Sladek, F. M., Darnell, J. E., Jr., & Crabtree, G. R. (1992) *Nature (London)* **355,** 457–461.
33. Pani, L., Overdier, D., Porcella, A., Qian, X., Lai, E. & Costa, R. H. (1992) *Mol. Cell. Biol.* **12,** 3723–3732.
34. Dirksen, M. L. & Jamrich, M. (1992) *Genes Dev.* **6,** 599–608.
35. Grossniklaus, U., Pearson, R. K. & Gehring, W. (1992) *Genes Dev.* **6,** 1030–1051.
36. Tao, W. & Lai, E. (1992) *Neuron* **8,** 957–966.
37. Li, C., Lai, C., Sigman, D. S. & Gaynor, R. B. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 7739–7743.
38. Li, C., Lusis, A. J., Sparkes, R., Tran, S. & Gaynor, R. (1992) *Genomics* **13,** 658–664.
39. Oliver, S. G., van der Aart, Q. J. M., Agostoni-Carbone, M. L., Aigle, M., Alberghina, L., *et al.* (1992) *Nature (London)* **357,** 38–46.
40. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Plainview, NY).
41. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
42. Hattori, M. & Sakaki, Y. (1986) *Anal. Biochem.* **152,** 232–238.
43. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 488–492.
44. Pani, L., Qian, X., Clevidence, D. & Costa, R. H. (1992) *Mol. Cell. Biol.* **12,** 552–562.
45. Tso, J. Y., Sun, X., Kao, T., Reece, K. S. & Wu, R. (1985) *Nucleic Acids Res.* **13,** 2485–2502.
46. Cao, Z., Umek, R. M. & McKnight, S. L. (1991) *Genes Dev.* **5,** 1538–1552.
47. Williams, S. C., Cantwell, C. A. & Johnson, P. F. (1991) *Genes Dev.* **5,** 1553–1567.
48. Lin, F. & Lane, M. D. (1992) *Genes Dev.* **6,** 533–544.
49. Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kishimoto, S., Nishio, Y., Nakajima, T., Hirano, T. & Kishimoto, T. (1990) *EMBO J.* **9,** 1897–1906.
50. Metz, R. & Ziff, E. (1991) *Genes Dev.* **5,** 1754–1766.
51. Wegner, M., Cao, Z. & Rosenfeld, M. G. (1992) *Science* **256,** 370–373.
52. Hacker, U., Grossniklaus, U., Gehring, W. J. & Jackle, H. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8754–8758.