A novel cDNA was partially isolated from a HepG2 cell expression library by screening with the promoter-linked coupling element (PCE), a site from the α-fetoprotein (AFP) gene promoter. The remainder of the cDNA was cloned from fetal liver RNA using random amplification of cDNA ends. The cDNA encodes a 239-amino acid peptide with domains closely related to the Drosophila factor nk-2. The new factor is the eighth vertebrate factor related to nk-2, hence nkx-2.8. Northern blot and reverse transcriptase polymerase chain reaction analysis demonstrated mRNA in HepG2, two other AFP-expressing human cell lines, and human fetal liver. Transcripts were not detected in adult liver. Cell-free translation produced DNA binding activity that gel shifts a PCE oligonucleotide. Cotransfection of nkx-2.8 expression and PCE reporter plasmids into HeLa cells demonstrated transcriptional activation; NH2-terminal deletion eliminated this activity. Cotransfection into AFP-producing hepatocytic cells repressed AFP reporter expression, suggesting that endogenous activity was already present in these cells. In contrast, cotransfection into an AFP-negative hepatocytic line produced moderate activation of the AFP gene. The cardiac developmental factor nkx-2.5 could substitute for nkx-2.8 in all transfection assays, whereas another related factor, thyroid transcription factor 1, showed a more limited range of substitution. Although the studies have yet to establish definitively that nkx-2.8 is the AFP gene regulator PCF, the two factors share a common DNA binding site, gel shift behavior, migration on SDS-acrylamide gels, and cellular distribution. Moreover, the nkx-2-related genes are developmental regulators, and nkx-2.8 is the first such factor associated with liver development.

Developmental processes are frequently associated with expression of specific homeobox transcription factors. Although all share a common form of DNA binding domain, the hundreds of known homeobox factors belong to many subfamilies with a wide variety of secondary domains (1). The Drosophila factor nk-2 is the prototype of a distinct family of homeobox factors. nk-2-related homeodomain factors have been characterized in Drosophila (nk-2, nk-3, bagpipe, and nk-4/tinman/msh2), plnarians (Dth1 and Dth2), leeches (lox10, Caenorhabditis (Ceh22), and vertebrates (nkx-2.1 to 2.7) (2). The nkx-2-related factors contain a characteristic secondary domain, the “conserved peptide,” which has an unknown function and is unrelated to known protein domains.

The three Drosophila homologues have important developmental functions. nk-2 is involved in early neurogenesis, nk-3 is required for visceral muscle formation, and nk-4 is essential for the formation of precardiac mesoderm.

The vertebrate nk-2 factors also regulate development. nkx-2.1, or thyroid transcription factor 1 (TTF-1), is a regulator of thyroid-specific gene expression, thyroid development, thyroid cell differentiation, and thyroid cell proliferation (3, 4) and is first expressed several days before thyroid differentiation (5). TTF-1/nkx-2.1 also regulates pulmonary development and gene expression (6, 7).

nkx-2.5/Csx and nkx-2.3 are both expressed in early cardiac primordia and thus replicate the function of tinman (8). nkx-2.6 is also expressed in heart. In Drosophila, ablation of tinman blocks cardiac development, whereas knockout of mouse nkx-2.5 arrests heart development at the looping stage (8). This is a less severe phenotype than the Drosophila knockout and probably reflects functional redundancy with nkx-2.3 and nkx-2.6 (9). nkx-2.7, another tinman homologue related to cardiac development, has recently been described in zebrafish (10).

nkx-2.2 is expressed in developing mouse brain, with an onset of expression at about 9 days gestation. The transcripts are found in localized regions that correspond to anatomic boundaries in the developing forebrain. The localization suggests that nkx-2.2 specifies differentiation of the developing diencephalon into its anatomically and functionally defined subregions (11).

In mice the cardiac mesenchyme forms from the ventral wall of the foregut at 8.5 days gestation, and the hepatic primordium buds from an adjacent area of the foregut at 8.5–9 days gestation (12). There is a strong association between this early phase of cardiac differentiation and nk-2-related factors, but no homeodomain factors of any sort have been associated with initial differentiation of the liver, later differentiation of bile ducts, or regulation of hepatic stem cells. Two homeobox factors, HNF-1 (13, 14) and HNF-6 (15), are associated with the

© 1998 by The American Society for Biochemistry and Molecular Biology, Inc.

Printed in U.S.A.

This paper is available on line at http://www.jbc.org
Homeobox Factor nkx-2.8

liver, but they regulate only the postdevelopmental phenotype. Another homeobox gene, hlx, has a limited effect on liver development. However, hlx is expressed in hematopoietic cells of fetal liver, not hepatocytes. Mice with targeted disruption of this gene undergo normal early hepatic differentiation, but the failure of hematopoietic cells to colonize the fetal liver results in a small but normally formed liver (16).

Our research has focused on the α-fetoprotein (AFP) gene, characteristically expressed from the earliest stages of liver development, but silent after birth. A site near the AFP promoter, the promoter-linked coupling element (PCE), appears to interact with the main developmental regulators of AFP expression. In HepG2 cells, the PCE binding activity has been characterized as a distinct transcription factor, PCF (17, 18).

To characterize PCF further, we used a PCE-containing oligonucleotide for expression cloning, leading to the isolation of nkx-2.8, which shares numerous properties with PCF. This is the first demonstration of an association between an nk-2 homeobox factor and liver gene expression.

EXPERIMENTAL PROCEDURES

Expression Cloning—A λgt11 library of oligo(dT)-primed cDNA from HepG2 cells (CLONTECH) was screened for plaques that bound the PCE in the AFP gene promoter (−166 to −155, TGGTCAAAGGACA; Ref. 18). This sequence and its complement were included in a double-stranded oligonucleotide with added TCGA sticky ends. Probe preparation and screening were modified from previously described methods (19, 20). The oligonucleotide was labeled with polynucleotide kinase and [γ-32P]ATP and concatenated with DNA ligase to an average size of 5–10 ng/μl. Bacteriophage plaques were plated using E. coli strain strain Y1090. Plaques were lifted onto Millipore HATF membranes, air dried, incubated in blocking buffer (5% nonfat dry milk, 25 mM HEPES, pH 7.9, 5 mM MgCl2, 40 mM KCl, 1 mM dithiothreitol) for 30 min and then probe mix (0.25% nonfat dry milk, 25 mM HEPES, pH 7.5, 5 mM MgCl2, 40 mM KCl, 0.1 mM dithiothreitol, 10 μg/ml sonicated denatured DNA, and 10 ng/ml labeled concatenated probe) for 12 h. DNA from Micrococcus lysodeiktikus was chosen for this mixture for its high G+C content (70%) and provided a very effective background competition that greatly enhanced the sensitivity of the screen. A screen of 1,000,000 phage plaques produced a single clone with appropriate binding specificity. A second screen of 1,000,000 plaques from another HepG2 library in pCR-Script SK+ (Stratagen, La Jolla, CA) provided no additional positives. Upon sequencing, the positive clone was found to be truncated at the 5′-end by an in-frame deletion that also removed part of the β-galactosidase gene. PCR assays (see below) were used to screen pooled DNA isolated from both libraries but detected only the single partially deleted clone. The cloned insert was subcloned in pBluescript KS+ (Stratagen). Because cloning instability was present and all further plasmids were propagated in a DNA repair and recombination-deficient E. coli host (SURE, Stratagen).

PCR-based Cloning—From the sequence of the new clone, direct and reverse transcriptase PCR assays (below) were used to screen a variety of RNA sources. A positive detection was obtained from a commercial cDNA prepared from human fetal liver (Marathon-Ready cDNA, CLONTECH) which was synthesized using random primers, cDNA ends cloning adaptor annealed to the 5′-end. Amplification was carried out with primers AP1 and R5A and Tha DNA polymerase (denaturation at 95 °C for 2 min; 7 cycles of 95 °C for 40 s, 72 °C for 4 min; 40 cycles of 95 °C for 40 s, 68 °C for 4 min). PCR primers

| Primer | Position | Sequence |
|--------|----------|----------|
| AP1    | 5′-Adaptor | CCACTCTAATACGACTCACTATAGGG |
| AP2    | 5′-Adaptor | CTACTATAGGGGACACCAATACG |
| F5A    | 601–628  | TCTAATGCAAGCTGAACGTACGTCC |
| F7     | 320–340  | GCCGCTGGTCGTTAGTCGGA |
| F9     | 542–562  | GAGACGTGGCGAGCACCTG |
| F10    | 246–370  | CCACCTACCTTCTTCGAGCAGAG |
| F12    | 5–25     | AGACGCCGACCTGGCGTTTC |
| R5     | 626–602  | AGACGCTTCAGTTGACAGTAT |
| R5A    | 628–597  | CGAGAGGCTTGCAGTTGACAGTAT |
| R10A   | 593–567  | CGAGATCTTCTGCAGTCC |
| R12A   | 376–346  | AGGCCTGTCTCCGCGAGAGG |
| R13A   | 342–320  | CTCGCGAATCCGGACAGCGGC |
| R16    | 1168–1141| AGGCGGACGCTAATAATCCTACTG |
| PBDG-A | 346–320  | CTGTGCTTGAAGCAGCAATGCG |
| PBDG-B | 346–320  | CAGGCAGCCTGTGCTTGAAG |

Primer positions for the nk-2.8 gene are numbered as in the sequence of Fig. 1.

| Primer | Position | Sequence |
|--------|----------|----------|
| AP1    | 5′-Adaptor | CCACTCTAATACGACTCACTATAGGG |
| AP2    | 5′-Adaptor | CTACTATAGGGGACACCAATACG |

For additional mapping and to rule out splicing variants the fetal liver cDNA and the gene were analyzed by direct PCR, using Tha DNA polymerase and a “touchdown” amplification protocol (denaturation at 95 °C for 2 min; 7 cycles of 95 °C for 40 s, 72 °C for 4 min; 40 cycles of 95 °C for 40 s, 68 °C for 4 min). mRNA Analysis—RNA isolation from cell lines and tissues, agarose gel electrophoresis, and Northern blot hybridization were carried out as described previously (22, 23). Poly(A) RNA was purified using a magnetic bead system (PolyATtract, Promega, Madison, WI). For each gel lane, 100 μg of total RNA was processed, 50 μg of E. coli tRNA was added and the cDNA was ethanol precipitated and redissolved for electrophoresis. For a hybridization probe, the F9:R5 PCR product was cloned in pCR-Script SK+. A riboprobe was generated with T7 RNA polymerase following plasmid linearization with PstI. For PCRs, 1 μl total RNA was incubated in a single tube with primers R5 and PBDG-B and Superscript II reverse transcriptase (Life Technologies) at 45 °C for 1 h. Amplification of nkx-2.8 transcripts was then carried out on an aliquot of the PCR mixture and R10A and KlenTaq DNA polymerase (CLONTECH). The sample was denatured at 95 °C for 2 min followed by 7 cycles of 95 °C for 40 s, 72 °C for 4 min; and 40 cycles of 95 °C for 40 s, 68 °C for 4 min. A control PCR for a housekeeping mRNA, porphobilinogen deaminase (PBDG; Ref. 24), was carried out with primers PBDG-A and PBDG-B and Taq DNA polymerase. The sample was denatured at 94 °C for 2 min followed by 40 cycles of 94 °C for 30 s, 55 °C for 20 s, and 72 °C for 30 s. Products were visualized on an ethidium bromide-stained agarose gel or transferred to nitrocellulose and detected by hybridization to a specific probe, the product of an F10:R7 PCR.

Cells, Tissues, and Transfection—Human cell lines HepG2, HuH7, and HuH1-clone 2 (Clone 2) are derived from hepatocellular carcinomas (25–27); and RPMI 7451 is from a cholangiocarcinoma (28). H4C3 is a rat hepatocellular carcinoma cell line that expresses high levels of albumin and low levels of AFP (23). Human lung carcinoma cell line H441 was used as a source of TTF-1 for gel shifts (29). All were propagated in Williams E medium containing 1–5% fetal calf serum, penicillin-streptomycin, and glucose. Specimens of human fetal and adult liver were provided by Dr. Stephen Strom (University of Pittsburgh).

CaPO4 transfection of cell lines HepG2, Clone 2, and HeLa was carried out as described previously (30). H4C3 cells were transfected with LipofectAMINE (Life Technologies, Inc.); 10-cm culture plates were inoculated with 1 × 106 cells and transfected after 2 days, using 10 μg of DNA and 60 μl of LipofectAMINE/plate, according to the supplier’s protocols. Individual transfection experiments always consisted of a...
**FIG. 1.** *nkx-2.8* cDNA sequence. The 5′-peptide and conserved peptide sequences common to *nk-2*-related factors are underlined, and the homeodomain is underlined and marked in bold type. Two short open reading frames in the 5′-untranslated region are marked with dotted lines. The residue at position 246 was C in 3 and G in 4 of the subclones that were sequenced, indicating a polymorphism.

Despite its shorter length, the *nkx-2.8* open reading frame defines a full-length factor. The 239-amino acid peptide of 25,810 Da is a basic protein with a predicted pl of 9.57. The cDNA also has an unusually high G+C content, 65.8% over its full length and 71.0% in the open reading frame.

Other regions of the *nkx-2.8* did not show significant relationships on BLAST analysis. However, FASTA alignment demonstrated significant homology to other *nk-2*-related factors. *nkx-2.8* has an NH2-terminal region comparable to *nk-2*-related factors, which have an 11-amino acid “5′-peptide” near the NH2 terminus (Fig. 2C). This peptide lies in a similar position in other family members except for human TTF-1, which has a unique 30-residue extension at the 5′-end. Although the 5′-peptide is not conserved in *nkx-2.8*, the nine residues beginning at amino acid 9 (TVRSLLGLP) show 44% identity to the 5′-peptide consensus. Four residues are identical, and the other five represent conservative substitutions.

**RESULTS**

cDNA Cloning and Sequence—Using a combination of approaches, overlapping products have been cloned which encode a novel transcription factor (Fig. 1). BLAST analysis of translated sequences showed two characteristic domains, a homeobox (Fig. 2A) and a conserved peptide (Fig. 2B). Both domains established a relationship to the *Drosophila* factor 9.2. These comparisons also demonstrated that the encoded factor is new, the shortest member of the family described so far. The new factor is the eighth vertebrate *nk-2*-related factor, hence *nkx-2.8*. 
with the NH$_2$-terminal region of human TTF-1 with about 30% identity. In addition, both TTF-1 and 
nkx-2.8 have polyglycine regions, although in slightly different locations. TTF-1 has a 
stretch of eight glycines between the homeobox and conserved 
peptides, whereas 
nkx-2.8 has six consecutive glycines just 
downstream of the conserved peptide.

**Gene Structure—**
The predicted cDNA structure was con-

firmed with a series of PCR studies of fetal liver cDNA, and 
parallel studies were carried out on genomic DNA. This anal-

ysis verified the unique cDNA and also localized an intron. The 
following primer combinations were studied: F7:R10A, F7:R7A, 
F7:R12A, F7:R13A, F12:R13A, and F5A:R16 (Table I). In the 
cDNA analysis, each combination gave a single strong product, 
suggesting that the mRNA has no splicing variants. Amplifi-
cation with F7:R7A and F7:R10A gave genomic DNA products 
750 bp larger than cDNA products, whereas genomic DNA 
did not amplify with F7:R12A. All other products were identical 
for both cDNA and genomic DNA. These results suggested a 
750-bp intron at the position of primer R12A. A genomic PCR 
product was cloned and partially sequenced (Fig. 3), confirming 
an intron within the codon located 31 amino acids from the 
NH$_2$-terminal side of the homeobox. This is similar to the 
structure of the human TTF-1 gene, which has a single 966-bp 
intron within a codon 36 residues proximal to the homeobox (29).

**mRNA Analysis—** Standard Northern blot analysis of total 
RNA was attempted with several probes, but because of the 
high G+C content, specific signals were obscured by nonspe-
cific hybridization to rRNA. To solve this problem, a riboprobe 
plasmid was constructed containing the F9:R5 amplimer, a 
region within the homeobox which had lower G+C content (Fig. 
4). In blots of total RNA, this probe detected a transcript with 
an apparent molecular mass of 1.6 kb but still showed cross-
hybridization to rRNA even under stringent conditions (not 
illustrated). However, when poly(A) RNA was isolated for blots, 
the hybridization was predominately to this single 1.6-kb tran-
script, which was clearly detected in RNA from three hepato-
carcinoma cell lines, HepG2, HuH7, and Clone 2 (Fig. 4B).

---

**Fig. 2. Comparison of nkx-2.8 with other nk-2-related factors.** Panel A, home-

odomains. Related homeodomains were aligned and listed above 
nkx-2.8 in order of decreasing similarity. A consen-
sus sequence was derived for nk-2-related factors, and a second consensus was de-

rived from 100 HOX genes closely related to the *Drosophila Antennapedia* gene. The 
nkx-2.8 sequence is in *bold* type; amino acids that deviate from the consen-
sus are shown in *light* type. The column labeled Δ shows the number of differences from 
the nk-2 consensus for each se-

quence. Predicted DNA base (:) and back-

bone (.) contacts are also marked (34). Panel B, conserved peptides. This nkx-2.8 
domain (*bold* type) is closely related to a consensus derived from other nk-2 fac-
tors. The domain consists of a nonhelical hydrophobic loop surrounded by highly 
charged, predominantly basic amino ac-
ids. The column labeled Δ shows the num-
ber of differences from the consensus. 
Panel C, 5′-peptides. Conserved peptides 
found in many nk-2-related factors near the 
5′-end are aligned. Position denotes the 
amino residue where each 5′-peptide is located. Conservative amino acid sub-
stitutions are *underlined*; nonconserva-
tive substitutions are shown in *lowercase*.

Panel D, comparison of nkx-2 proteins. 
nkx-2.8 is shorter than the related factors 
but aligns at the 5′-ends of nkx-2.2, nkx-

2.5, and murine TTF-1, and with amino 
acid 30 of human TTF-1, the longest pro-
tein in the family. All align at the 3′-end. 
Studies of protein functional regions, re-
ported for murine nkx-2.5 (33) and human 
TTF-1 (4), are also summarized.

---

**Homeobox Factor nkx-2.8**

---

**Conserved Peptides**

---

**Proteins**

---

**5′ Peptides**

---
FIG. 3. Gene sequence. A genomic PCR product was cloned and sequenced to define the junctions of the single intron. The sequence of this genomic is aligned with the cDNA sequence and with consensus splicing motifs (35).

FIG. 4. mRNA analysis. Panel A, probe and PCR strategies. The map shows the exon structure of nkx-2.8. The shaded area represents the homeodomain. The PouII site was used to linearize a plasmid for generation of an antisense riboprobe. PCR primers are shown as arrows. Panel B, Northern blots of poly(A) RNA. Left, a series of RNAs from cell lines and adult human tissues: hepatocellular carcinoma cell lines HepG2, HuH7, and Clone 2; cholangiocarcinoma cell line RPMI 7451; and adult liver, spleen, and lung. Right, a separate hybridization study compares three human fetal liver specimens (1, 18 weeks gestation; 2, 18 weeks; 3, 21 weeks) and three adult liver specimens (1, 63 years; 2, 57 years; 3, 16 years). Each lane in both panels contains the poly(A) RNA purified from 100 μg total RNA. Panel C, reverse transcriptase PCR analysis. 1-μg aliquots of total RNA were treated with reverse transcriptase and primers from both the fetal liver mRNAs. The Northern blot studies can only be considered semiquantitative because poly(A) RNA preparations vary significantly.

Reverse transcriptase PCR analysis was also set up (Fig. 4C) because the Northern blot detection in fetal liver was inconclusive, and the possibility of even lower transcript levels in adult liver could not be ruled out. Reverse transcriptase PCR demonstrated a 273-bp product predicted for spliced transcripts in two cell lines and fetal liver. Hybridization confirmed the identity of the PCR product and indicated that larger products detected in some of the RNA preparations were nonspecific. On long film exposures, a very weak signal was also detected in lung RNA, but the analysis of adult liver was completely negative. Thus nkx-2.8 mRNA was detected unequivocally in fetal but not adult liver and in three cell lines with phenotypes that resemble fetal liver.

Analysis of Translated nkx-2.8 Protein—The full-length and deleted nkx-2.8 expression plasmids were translated in vitro and the products analyzed by oligonucleotide gel shift. Gel shifts (Fig. 5A) were carried out with a PCE oligonucleotide (tgaTGTTCAAGGACA). Full-length nkx-2.8 protein produced a band shift that comigrated with the PCE band shift of HepG2 cells but was distinctly lower than the band shift of the related factor, TTF-1. As expected, the 5′-deleted protein produced a band shift lower than that of intact nkx-2.8. The analysis demonstrated that the expression plasmids containing the cDNAs are functional and encode peptides that bind the PCE.

The gel shift analysis showed a similarity between PCF and nkx-2.8. However, the latter had a predicted molecular mass of 26 kDa, whereas two types of experiments indicated a higher molecular mass for PCF as measured on SDS-acrylamide gels: UV-cross-linking indicated peptides of about 34 kDa, whereas partial purification enriched three bands of 32–34 kDa (18). Translated protein migrated as a single band in SDS-acrylamide electrophoresis, with an apparent molecular mass of 33.6 kDa (Fig. 5B). Basic proteins like nkx-2.8 frequently show anomalous slow migration on SDS-acrylamide gels. The observed migration is common to both nkx-2.8 and PCF.

nkx-2.8 Function in a Nonhepatocytic Cell Line—HeLa cells, expected to be free of endogenous expression of nkx-2.8 and other hepatocytic transcription factors, were used for the initial evaluation of transcriptional activity. nkx-2.8 was analyzed directly and compared with the related factors nkx-2.5 and TTF-1 (Fig. 6). For this analysis, a reporter plasmid, pCE4-HIV-CAT, was constructed in which four PCE sites were placed in tandem near a basal promoter. In HeLa cells, the reporter showed clear activation over unstimulated levels. Maximum 9-fold stimulation occurred at 1 μg. Higher levels had somewhat lower activity but still stimulated the reporter. nkx-2.5 also showed comparable (14-fold) activation. The deleted form of nkx-2.8 showed no transcriptional stimulation, nor did TTF-1, even though the latter showed clear binding to a PCF oligonucleotide (Fig. 5). Thus nkx-2.8 is a functional transcriptional activator, can act through a PCE site, and requires the intact homeodomain. The full-length and 5′-deleted protein produced a band shift lower than that of intact nkx-2.8. The analysis demonstrated that the expression plasmids containing the cDNAs are functional and encode peptides that bind the PCE.

nkx-2.8 Function in Hepatocytic Cell Lines—For evaluation of nkx-2.8 in hepatocytic cells, we attempted to study differentiated hepatocytic cell lines with both fetal (AFP + , albumin + ) and adult (AFP − , albumin − ) phenotypes (Fig. 7A). In general, hepatocellular carcinoma cell lines either have a fetal phenotype like HepG2 or are undifferentiated. Cell lines reported to have an adult hepatocytic phenotype are infrequent and problematic, although we utilized two such lines, Clone 2 and H4C3. Morigawa et al. (27) characterized gene expression in the HuH1 (Clone 2). HuH1 has a fetal phenotype, but Clone

2 N. Crawford and J. Locker, unpublished results.
2 was reported to have selectively lost AFP expression. However, our analysis of Clone 2 showed reduced but still significant AFP expression. Moreover, we also observed clear nkx-2.8 expression in Clone 2 (Fig. 4). Reuber hepatoma cell lines like H4C3, common models for the adult hepatocyte phenotype, also have problematic features. Because H4C3 is rat-derived, we did not analyze for nkx-2.8 transcripts. However, we previously detected low AFP mRNA expression (23) and a weak PCF gel shift from these cells (18).2

A variety of reporter genes were evaluated in the three hepatocytic cell lines. To indicate the cell differentiation states in these lines, the relative activities of the reporters are compared in Fig. 7A. Fig. 7B illustrates two experiments in HepG2 cells. The data on the left show the responses of two reporters to increasing levels of transfected nkx-2.8. The data on the right compare several nk-2 family expression plasmids transfected at a single DNA concentration. In HepG2 cells, the reporter pPCE4-HIV-CAT was stimulated strongly (14-fold) compared with the same promoter without the PCE binding site, indicating a significant level of an endogenous activating factor binding at the PCE. The transfected expression plasmids all reduced expression of this reporter gene; they also repressed a very active reporter containing intact AFP gene transcription controls. The repression was relatively weak and observed with intact and deleted nkx-2.8, TTF-1, and nkx-2.5, although these may not all repress by the same mechanism. Comparison of the two illustrated nkx-2.8 analyses shows the general reproducibility of transfections in separate experiments; the levels were 76 ± 19% and 51 ± 7% of the control (reporter plasmid alone).

Most likely, repression in HepG2 resulted from squelching, i.e. overexpression of transcription factor and competition for trans-activators or binding sites for protein-protein interaction. Such squelching is probably apparent in the HeLa cell studies, where greater amounts of transfected plasmid produced less stimulation than 1 µg. Thus the repression in HepG2 cells may have occurred because HepG2 already contained significant levels of a PCE-activating factor. An alternate possibility is competition for binding sites in the absence of transcriptional activation. This is most likely true for the deleted nkx-2.8, which lacks an activation domain, but nkx-2.8, nkx-2.5, and TTF-1 are clearly activators in other settings (e.g. HeLa and H4C3 cells). Nevertheless, there is a formal possibility that the factors cannot interact with the promoters of either the PCE reporter or the AFP reporter in the transcriptional environment of HepG2 cells. In this case, it would have to be presumed that another trans-factor activates in these cells through binding at the PCE and that all three factors, nkx-2.8, nkx-2.5, and TTF compete for binding of this factor but cannot activate.

Transfection of both intact and deleted nkx-2.8 also repressed the AFP gene reporter in Clone 2 cells (Fig. 7C), although in these cells, the activity of the PCE reporter was too low to show significant effects. The fact that both HepG2 and Clone 2 showed repression by a PCE-targeted factor suggests that endogenous AFP gene controls active at the PCE were intact in both cell lines. Clone 2 has much less AFP gene expression than its parental line, but this apparently reflects altered regulation at gene sites other than the PCE and probably does not represent a true adult hepatocytic phenotype.

In contrast to the other hepatocytic lines, H4C3 cells (Fig.

![Fig. 5. Analysis of in vitro translated nkx-2.8. Panel A, oligonucleotide gel shift analysis. Gel shift was carried out using a double-stranded end-labeled PCE oligonucleotide, tcgaTGTTCAAGGACA. The lanes contained 6 µl of in vitro translation mixture programmed with pCMV-Nkx2.8 (labeled Nkx2.8), pCMV-Nkx2.8Δ (Deletion), or a luciferase-encoding plasmid (Control). Two additional lanes show the same oligonucleotide shifted with HepG2 and H441 cell extracts to demonstrate the positions of PCF and TTF-1 gel shifts. Panel B, SDS-acrylamide gel electrophoresis of labeled nkx-2.8. An in vitro translation including [35S]leucine was programmed either with pCMV-Nkx2.8 (Nkx2.8) or without plasmid DNA (Control). The labeled product was resolved on a 10% acrylamide and SDS gel. The figure is a composite showing autoradiograms detected with a PhosphorImager and Coomassie Blue-stained markers.](http://www.jbc.org/)

![Fig. 6. Transient expression in HeLa cells. Transfections were carried out with 15 µg of total DNA, including 10 of the reporter plasmid pPCE4-HIV-CAT, an expression plasmid (pCMV-Nkx2.8 (Nkx2.8) at 0, 0.5, 1, 2, and 5 µg; 2 µg of pCMV-Nkx2.8Δ (Deleted), pCMV-TTF1 (TTF), or pCGN-Nkx2.5 (Nkx2.5), and carrier DNA. The data are shown as average values and standard deviations of each group, and all are compared with the expression of the nkx-2.8 reporter in the absence of expression plasmids (100%).](http://www.jbc.org/)
7D) showed some stimulation of both AFP and PCE reporters. Intact nkx-2.8 showed only weak effects. Although there was no significant stimulation of the PCE reporter in two separate experiments, an AFP gene reporter was stimulated 22% in the illustrated experiment and 20% in a separate experiment (not illustrated). The deleted nkx-2.8 had no effect on the reporters, whereas nkx-2.5 and TTF-1 showed greater than 3-fold stimulation in some combinations. The transfections with the latter two factors demonstrate that nk-2-related factors, acting through the PCE, are potentially strong AFP activators in H4C3 cells.

Notably, nkx-2.8 did not induce AFP expression in H4C3 cells up to the levels of HepG2 cells. However, such a result would be expected only if lack of nkx-2.8 expression was the single property that prevented strong AFP gene expression in H4C3 cells. Interestingly, TTF-1 activated the reporters in H4C3 but not in HeLa, suggesting that a variety of cofactors with distinct cellular distributions interact with various nk-2-related factors.

**DISCUSSION**

An Unusual Gene and Transcript—nkx-2.8 mRNA has an unusual DNA sequence. The overall G+C content is 66%, whereas the coding region is 71%. Moreover, the exons contain 129 CG dinucleotides, potential DNA methylation sites. Such CG islands are found upstream of some genes but are rarely incorporated into coding regions. A high G+C content is also a feature of the mammalian TTF-1 (71% G+C), nkx-2.2 (65%), and nkx-2.5 (70%) genes (11, 29, 36). Interestingly, when examples of these genes have been sequenced from lower vertebrates, the high G+C content is not conserved (10, 37). The *Xenopus* nkx-2.5 gene, for example, is only 48% G+C.

Computer analysis of RNA folding predicts an unusually high degree of secondary structure for the nkx-2.8 mRNA. The combination of high G+C content and secondary structure has led to technical problems in the cloning, purification, and quantification of the mRNA. The abundance of nkx-2.8 mRNA may have been underestimated because of these technical problems.
TTF1  
GNNACCTCAAG

PCF  
WGTTCAGGACA

Consensus 1  
TCAAG

Consensus 2  
RYTCAAG

**Fig. 8.** DNA binding sites. The PCF binding site (PCE) has been characterized at ~166 to ~155 in the AFP gene promoter (18). For alignment, this is shown as the reverse strand sequence. The TTF-1 site is a consensus compiled from eight binding sites by Guazzi et al. (3). The sites share a common motif (TCAAG) and do not show the ATTA motif typical of Antennapedia class homeobox binding sites.

Nevertheless, a low level of mRNA is characteristic of the nk-2 gene family except for TTF-1 and may typify genes that have limited regulatory function. The studies of this paper provide a technical paradigm for a difficult class of mRNAs.

Relationship of Binding Sites—It is not surprising that there is a close relationship between the PCE and TTF-1 binding sites (Fig. 8) because the nk-2.8 homebox is extremely similar to the TTF-1 homebox and to those of other members of the family. The sites share a common motif, TCAAG, which is also the central motif of the Drosophila nk-2 binding site (38). nkx-2.5 also binds strongly to a TTF-1 site (33). The structure and sequence of these nk-2-family homeodomains differ considerably from the more common Antennapedia class homeodomains. The latter bind to a TAAT central motif that corresponds to CAAG in the nk-2 family sites (39).

**nkx-2.8 and PCF—**nkx-2.8 was obtained through efforts to clone the AFP gene developmental regulator activity we characterized as PCF (18). There are marked similarities between PCF and nkx-2.8. Both bind the PCE and activate from that site, and both are expressed in the same cells, along with AFP. However, the exact peptide composition of PCF is not fully resolved. Photoaffinity labeling demonstrated a broad PCF band migrating at 34 kDa, and partially purified PCF showed multiple bands in the 32–34-kDa range. nkx-2.8 is basic, and PCF also purifies as a basic protein(s). Although only 26 kDa, nkx-2.8 has anomalous migration on SDS-acrylamide gels which causes it to migrate near 34 kDa, the position of PCF components. The multiple PCF bands might represent isoforms, post-translational modifications, related factors, coactivators, or even unrelated peptides that copurify. It appears likely, however, that nkx-2.8 is a component of PCF.

There is a more significant problem in the relationship between PCF and nkx-2.8. Surprisingly, transfection of nkx-2.8 into AFP-positive lines HepG2 and Clone 2 caused transcriptional repression of AFP gene reporters, whereas transfection into H4C3 cells led only to weak AFP gene activation. In contrast, PCF has been characterized as a strong activator in HepG2 in H4C3 cells led only to weak AFP gene activation. Moreover, the levels of FTF are highest in adult liver, where the AFP gene is silent. Recent studies in Drosophila have shown an important relationship that may explain interaction of PCF, nkx-2.8, and FTF. In Drosophila, FTZ-TF1 and the homeodomain protein ftz are mutually dependent cofactors. The two factors interact to activate important low affinity sites (43, 44). Similar interaction between nkx-2.8 and FTF, or other members of each family, might account for the complex regulation that is observed at the PCE.

**nkx-2.8 and Other nk-2-related Factors—**The nk-2-related factors are defined by distinctive homeboxes, and the studies in this paper demonstrate that nkx-2.8, nkx-2.5, and TTF-1 all bind the PCE. nkx-2.5 substituted for nkx-2.8 in all of the transfection assays, suggesting very similar function for these two factors, whereas TTF-1 showed more limited ability to substitute. Like many activation domains, the nkx-2.8NH₂ terminus is rich in proline and acid residues, and the deleted 31-amino acid region shares this composition. Direct NH₂-terminal deletion eliminated activation function. In contrast, nkx-2.5 and TTF-1 have terminal regulatory domains with more internal activation domains because short NH₂-terminal deletions increased transcriptional activation before more extensive NH₂-terminal deletion removed the apparent activation domain (4, 33). nkx-2.8 has a shorter NH₂-terminal region (86 amino acids) than either nkx2.5 (136 amino acids) or murine TTF-1 (159 amino acids). nkx-2.8 is thus more compact than other members of the family and may lack an NH₂-terminal regulatory region, but it nevertheless retains activation function in its shortened NH₂ terminus. Moreover, the unique NH₂-terminal domain of nkx-2.8 might be subject to specific regulation in hepatocytic cells which is evaded by nkx-2.5 and TTF-1.

The present studies have established that nkx-2.8 gene expression is associated with AFP expression in fetal but not adult liver and in hepatocellular carcinoma. AFP is expressed not only in fetal liver but also in primitive endoderm and yolk sac. A common target in the AFP gene could be regulated by different members of the nk-2 gene family in these settings, just as cardiac development involves the partially redundant expression of several different nk-2-related genes (2, 8–10). Future areas of investigation include the role of nkx-2.8 overexpression in neoplasia and the identification of additional nkx-2 genes related to endodermal development.

Our findings that associate nkx-2.8 with liver development remain preliminary, limited by our initial choice of a human cell model, because the human is not a suitable organism for the detailed study of development. For characterization of animal development, we recently cloned segments of the rat and mouse nkx-2.8 genes and found that they are extremely similar to their human counterpart in both NH₂-terminal and homeobox regions.3 Such strong conservation suggests a gene with important function.

Acknowledgments—We are grateful to Steven Strom for providing specimens of human fetal and adult liver and to Roberto Di Lauro and Robert J. Schwartz for providing TTF-1 and nkx-2.8 expression plasmids.

REFERENCES
1. Duboule, D. (1994) Guidebook to the Homeobox Genes, Oxford University Press, Oxford
2. Durocher, D., Chen, C.-Y., Ardati, A., Schwartz, R. J., and Nemer, M. (1996) Mol. Cell. Biol. 16, 4648–4655
3. Guazzi, S., Price, M., DeFelice, M., Damante, G., Mattei, M.-G., and DiLauro, G. Apergis and J. Locker, unpublished results.
A Novel nk-2-related Transcription Factor Associated with Human Fetal Liver and Hepatocellular Carcinoma

George A. Apergis, Nancy Crawford, David Ghosh, Claire M. Steppan, William R. Vorachek, Ping Wen and Joseph Locker

J. Biol. Chem. 1998, 273:2917-2925.
doi: 10.1074/jbc.273.5.2917

Access the most updated version of this article at http://www.jbc.org/content/273/5/2917

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 42 references, 19 of which can be accessed free at http://www.jbc.org/content/273/5/2917.full.html#ref-list-1