HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally

Eseng Lai,1 Vincent R. Prezioso, Eleanor Smith, Oksana Litvin, Robert H. Costa,2 and James E. Darnell, Jr.3
Laboratory of Molecular Cell Biology, Rockefeller University, New York, New York 10021 USA

Hepatocyte-specific gene expression requires the interaction of many proteins with multiple binding sites in the regulatory regions. HNF-3 is a site found to be important in the maximal hepatocyte-specific expression of several genes. We find that liver nuclear extracts contain three major binding activities for this site, which we call HNF-3A, HNF-3B, and HNF-3C. Purification from rat liver nuclear extracts of HNF-3A and HNF-3C reveals that each activity corresponds to a distinct polypeptide, as determined by SDS-PAGE. Peptide sequence derived from the most abundant species, HNF-3A, was used for synthesizing probes with which to isolate a cDNA clone of this protein. The encoded protein contains 466 amino acids (48.7 kD) and has binding properties identical to those of the purified protein. A 160-amino-acid region that does not resemble the binding domain of any known transcription factor is essential for DNA binding. The mRNA for HNF-3A is present in the rat liver but not in brain, kidney, intestine, or spleen, and the basis for this difference is cell-specific regulation of HNF-3A gene transcription.

[Key Words: Transcriptional control; hepatocyte-specific factor; HNF-3]

Received May 4, 1990; revised version accepted June 8, 1990.

Cell-specific transcriptional control exists for many specialized products of cells in all multicellular organisms. It is also the basis of the limited differences in cell type among single-cell eukaryotes (e.g., a and α yeast cell types). In this latter case, some of the transcription factors that act in cell-type-specific transcription exist in both cell types, but others are transcriptionally regulated factors that actually determine cell type (Nasmyth and Shore 1987; Herskowitz 1989). How do we decide the basis for the demonstrated transcriptional control (Derman et al. 1981; Powell et al. 1984) of a whole series of genes in a mammalian tissue such as the liver? Does regulated transcription in adult hepatocytes rely on one or several transcription factors that are themselves limited to hepatocytes, or are there generally distributed factors that are specially modified in the hepatocyte? To answer these questions, the identification of DNA-binding proteins that bind to sites shown to be required for hepatoma-specific expression of genes has been a fruitful initial approach (Costa et al. 1986, 1988; Cerghini et al. 1987; Courtois et al. 1987; Lichtsteiner et al. 1987; Monaci et al. 1988). The field has progressed to the identification of at least five nuclear protein factors detected by site-specific DNA-binding assays that appear to be limited to or at least more abundant in liver extracts than in extracts of other cell types. To know whether these factors are products of independent genes and are themselves regulated transcriptionally, it is necessary to clone the genes encoding these factors and determine their transcription patterns. The genes for two of these proteins have been characterized: C/EBP, the original leucine zipper protein (Landschulz et al. 1988) and HNF-1 [LF-B1], a protein with distant similarity to the POU homeo domain proteins (Frain et al. 1989; Baumhueter et al. 1990; Finney 1990). In this paper we report the cloning of the cDNA of a third liver-enriched transcription factor called HNF-3A. This protein may have a more limited tissue distribution than the other two factors reported previously, and the limited distribution appears to be based on transcriptional control of the HNF-3A gene. From sequence analysis and comparison with other known protein structures, the HNF-3A protein contains a DNA-binding structure unlike that of any other DNA-binding protein.

We identified this protein originally in rat liver nuclear extracts, because it binds specifically to sequences proved to be functionally important in the hepatocyte-specific expression of two genes from the mouse, TTR [transthyretin] and α-1 antitrypsin (Costa et al. 1989).
Mutation of the HNF-3-binding site in an otherwise intact TTR promoter reduced expression of the TTR genes in HepG2 cells dramatically. Furthermore, a multimer of the HNF-3-binding site was sufficient to stimulate expression of a heterologous promoter in HepG2 cells but not in HeLa cells. The HNF-3-binding activity, itself, was highly enriched in liver cell nuclear extracts relative to those of other tissues, including spleen, kidney, brain, and intestine. All of these results suggested that proteins interacting with the HNF-3 site are essential for optimal expression of the mouse TTR and α-1-antitrypsin genes in the liver and made any such proteins good candidates for study as a possible transcriptionally controlled transcription factor.

**Results**

Purification and protein sequencing of HNF-3A

HNF-3-binding activity was monitored by use of a double-stranded oligonucleotide from the mouse TTR promoter (−111 to −85). In crude liver nuclear extracts, there appear to be three specific protein–DNA complexes detectable by gel mobility-shift assays, whereas nuclear extracts of kidney, brain, and spleen fail to form any specific complexes [Costa et al. 1989]. The liver proteins responsible for the complexes can be separated partially by chromatography on heparin–agarose (Fig. 1). It is not known whether these different complexes represent binding by different modifications or proteolytic fragments of a single protein or by different gene products. The results described below suggest that each complex represents site-specific binding to the HNF-3 oligonucleotide of different-sized molecules rather than of multimers of a single protein. Thus, we will refer to the molecules responsible for complexes A, B, and C as HNF-3A, HNF-3B, and HNF-3C, respectively. In our initial characterization of the HNF-3-binding activity, we separated complexes A and B from complex C to learn something about the basis of each complex and then proceeded to focus our efforts on HNF-3A.

Purification of HNF-3A to near homogeneity was achieved through five separate columns ending with three passes through the sequence-specific DNA affinity column (Table 1). HNF-3B is also present in these preparations. Figure 2A shows an SDS-polyacrylamide gel of the proteins present through the last steps of the purification. The major band (lane 6) migrated at an apparent molecular mass of 50,000 daltons. The verification of this band as HNF-3A is shown in Figure 2, B and C. Proteins in slices of a parallel SDS-polyacrylamide gel containing the identical affinity-purified material to that shown in Figure 2A, lane 6, were extracted, renatured, and assayed for DNA-binding activity. The renatured protein from slice 3, which contained the 50-kD band, bound specifically to the HNF-3 oligomer and yielded a gel mobility-shift complex that is identical in mobility to that of HNF-3A in liver nuclear extracts (Fig. 2C). The smaller amount of protein of ~46 kD in slice 2 yielded a faster migrating shift complex corresponding to HNF-3B (Fig. 2C). It is difficult to distinguish between a degradation product of HNF-3A or HNF-3B and a trace amount of HNF-3C protein, which may also be present. The major contaminants at this stage, present in slice 4, do not possess renaturable DNA-binding activity. The final affinity-purified material shown in Figure 2B (which represents a sample used for amino acid sequence determination described later) contained a 50-kD protein of HNF-3A and a small amount of a 46-kD protein corresponding to HNF-3B.

A similar but shorter purification scheme was used for HNF-3C, which yielded a major protein of 43 kD (Fig. 3A). This protein was confirmed by renaturation studies similar to those described above to be responsible for the C complex (Fig. 3B). A small amount of HNF-3A was present in the HNF-3C preparation and formed a slower migrating complex (Fig. 3B, slice 3). Thus, HNF-3A, HNF-3B, and HNF-3C each formed sequence-specific DNA–protein complexes that migrated successively faster in parallel with the decreasing polypeptide sizes of the three proteins.

Table 1 summarizes the results of each step in the purification of HNF-3A and HNF-3C. A total of 3 μg of HNF-3A and 1.3 μg of HNF-3C was obtained from 250 mg of nuclear extract. (NP-40 was added to the fractions eluted from the Sephacyr-S300 to stabilize the binding activity.) The DNA–cellulose column included in the scheme removed nuclease activity that otherwise would have destroyed the specific DNA affinity column. The preparation of HNF-3A after column chromatography

---

**Figure 1.** Three HNF-3 protein DNA complexes. A liver nuclear extract was chromatographed on a heparin–agarose column. Fractions eluted with a linear NaCl gradient were assayed by gel mobility shift with the strong HNF-3-binding site from the transthyretin gene as labeled probe (−85 to −111). Approximate NaCl concentrations based on conductivity measurements were 0.2, 0.32, and 0.45 M for fractions 10, 30, and 50, respectively. Proteins in liver nuclear extract form multiple gel mobility-shift complexes. The major complexes are designated A, B, and C. Free probe is labeled F.
Table 1. Purification summary of HNF-3A and HNF-3C

|                | Total protein | Activity | Purification | Yield |
|----------------|---------------|----------|--------------|-------|
|                | [mg]          | [units]  | (fold)       | (%)   |
| HNF-3A         |               |          |              |       |
| nuclear extract| 250           | 20,000   | 100          |       |
| heparin–agarose| 50            | 16,000   | 4            | 80    |
| Sephacryl S-300| 5             | 13,000   | 32           | 65    |
| FPLC Mono S    | 3             | 10,000   | 42           | 50    |
| DNA–cellulose  | 2             | 8000     | 50           | 40    |
| HNF-3 affinity x 3 | 0.003       | 6000     | 25,000       | 30    |
| HNF-3C         |               |          |              |       |
| Nuclear extract| 250           | 8000     |              |       |
| heparin–agarose| 24            | 7200     | 9            | 90    |
| Sephacryl S-300| 2.5           | 6000     | 72           | 75    |
| FPLC Mono S    | 1.6           | 5000     | 94           | 63    |
| DNA–cellulose  | 1             | 4200     | 133          | 53    |
| HNF-3 affinity x 2 | 0.0013       | 2800     | 69,200       | 36    |

Protein was quantitated by using a colorometric method (Bio-Rad) with BSA as standard. After the affinity step, quantitation was by silver and Coomassie blue stain. DNA-binding activity was determined with gel mobility-shift assays. One unit is defined as the activity that binds 0.1 ng of probe under standard assay conditions. Because the binding is performed in the presence of nonspecific competitor prior to the affinity column, the quantitation is not directly comparable before and after the affinity column. As defined, the values probably underestimate activity in the nuclear extract and overestimate both the final yield and fold purification.

was estimated to be ~75% pure (see Fig. 2B), whereas the sample of HNF-3C was ~90% pure. To obtain pure protein for sequencing, each sample was applied to an SDS-polyacrylamide gel and electroblotted to a PVDF membrane. The membrane was stained with Coomassie Blue, and either the 50-kD or the 43-kD band was excised and submitted for sequence analysis. Two attempts for each preparation yielded no sequence information, presumably due to blockage of the amino terminus of each protein.

At this point, we decided to focus on the more abundant and larger HNF-3A. We sought to generate and purify peptide fragments of HNF-3A for sequencing. Ten micrograms of purified protein was obtained from 800 mg of nuclear extract after column chromatography. An SDS-polyacrylamide gel was again used to isolate the 50-kD protein. After electroelution and removal of SDS, the protein was digested with endoproteinase lys-C, which cleaves specifically at lysine residues. The peptides were then separated by reverse-phase HPLC on a C-18 column. Four peaks were submitted for sequencing. Two peptide samples yielded unambiguous sequences, which allowed the design of oligonucleotide probes 157 and 144 for screening a rat liver cDNA library (Fig. 4).

Cloning and sequencing of HNF-3A

From 1,000,000 λ phage recombinants, ~50 positive clones were isolated with probe 144. Because this probe contained inosine residues, we expected to identify false-positive clones. Therefore, we screened these clones with probe 157 and found that only one hybridized to both probes. Sequence of this clone showed that it encoded both peptides derived from the purified protein. Subsequently, this clone was used to screen at high stringency another 1,000,000 recombinants from the same library. An additional nine independent clones were identified. The longest recombinant, H3, contained a 2.2-kb insert and was used for most of the subsequent analyses. The sequence of H3, shown in Figure 5, contains the entire coding sequence of HNF-3A but is missing a short stretch of the 3'-untranslated region. An open reading frame encoding a protein of 466 amino acids begins at base 58 preceded by a sequence that fits well with the consensus for translation initiation [Kozak 1986]. The calculated molecular weight of 48,700 daltons agrees well with the molecular mass estimate of 50 kD based on SDS-polyacrylamide gel analysis of the purified protein, and the protein product derived from this cDNA by in vitro transcription and translation has a mobility on SDS-polyacrylamide gels identical to that of the purified protein.

Because the sequence contains multiple methionines at the amino-terminal end, we determined whether the full-length protein resulted from translation initiation at either the first or second methionine. RNA for translation was produced by SP6 or T7 polymerase by copying the entire H3 sequence. Reticulocyte extracts produced the full-length 50-kD band from the full-length RNA (Fig. 6A, lane 1). A deletion construct lacking the first 98 nucleotides, just beyond the second in-frame AUG, was then used to generate RNA for in vitro translation. The translation product from the deletion (shown in lane 2) was slightly but distinctly smaller than the full-length protein. Therefore, the preferred initiator AUG appeared to be removed by the deletion, resulting in initiation at a downstream site in the reticulocyte lysate.

DNA binding of in vitro-translated HNF-3A and delineation of residues required for binding

Verification that the H3 clone encoded the HNF-3A DNA-binding activity was obtained in several ways.
Figure 2. Renaturation of HNF-3A activity. (A) A silver-stained SDS-polyacrylamide gel of samples from the later stages of the purification of HNF-3A. (Lane 1) Mono S pool (Mono S); (lane 2) first affinity pool (Aff 1); (lane 3) second affinity flowthrough (FT Aff 2); (lane 4) molecular weight markers (MWM); (lane 5) second affinity column, fraction 2 (Fx 2 Aff 2); (lane 6) second affinity column, fraction 8 (Fx 8 Aff 2), which contained the peak binding activity. Arrows indicate the position of cuts made in a parallel gel of fraction 8 to yield six gel slices. Protein from these gel slices was eluted, renatured, and assayed for binding activity, as shown in C. (B) A sample of the material pooled from a third-pass HNF-3 affinity column showing the major 50-kD protein band and a minor 46-kD protein. (C) Autoradiograph of a gel mobility-shift assay with renatured samples from gel slices 1–6. (Lane 7) Total sample from affinity fraction 8, diluted 1:50. Renaturation recovery is estimated at ~20%. Slice 3, corresponding to the 50-kD protein region, yielded a complex identical to complex A, the major complex of the starting sample. Slice 2 contained a 46-kD protein that yielded a faster migrating complex equivalent to HNF-3B in Fig. 1. Slice 2 also contained a trace of activity corresponding to complex C.

First, the sequence of H3 contains both sequenced peptides derived from purified HNF-3A protein [Fig. 5]. Second, we demonstrated that the protein product of H3 binds specifically to the HNF-3 site of the TTR promoter. In vitro protein products translated from SP6 or T7 RNA polymerase transcripts bound to the HNF-3-binding site, as tested by gel mobility-retardation assays with a double-stranded oligomer of the HNF-3-binding site (−107 to −92; Fig. 6B). The specific DNA–protein complex formed with in vitro-produced protein migrated at the same rate as that of the complex with partially purified rat HNF-3A protein. The in vitro-translated product was prevented from forming a complex only by an oligonucleotide representing the HNF-3-binding site [Fig. 6B, lane 2]. The identity of in vitro-translated product of the H3 clone with the liver protein HNF-3A was demonstrated further by a footprint assay [Fig. 6C]. The binding of the protein product of the H3 clone to a fragment of the TTR promoter protected the nucleotides

Figure 3. Renaturation of HNF-3C activity. (A) A silver-stained gel containing samples of complex C [see Fig. 1], bound and eluted twice from a specific oligonucleotide affinity column. (Lane 1) Starting material loaded onto column (START); (lane 2) flowthrough fraction (FT); (lane 3) 0.15 M NaCl wash (Wash); (lane 4) markers (M); (lane 5) fraction 4 (Fx 4); (lane 6) fraction 8 (Fx 8); (lane 7) fraction 12 (Fx 12). Fraction 8 contains the peak binding activity. Arrows indicate the position of cuts made to a parallel gel lane with fraction 8 to yield six gel slices. Protein from these gel slices was eluted, renatured, and assayed for binding activity, as shown in B. (B) An autoradiograph of a gel mobility-shift assay with renatured samples from gel slices 1–6. (Lane 7) Total sample from fraction 8, diluted 1:50. Slice 3 contained the major 43-kD protein and yielded a complex identical to complex C, whereas slice 4 contained a small amount of protein yielding complex A. (50K) Sample from fraction 8 of HNF-3A (see Fig. 2), which yielded predominantly complex A.
transcribed from -97 to -107 upstream of the cap site from chemical cleavage (Fig. 6C). This footprint exactly matches that of the partially purified rat liver HNF-3 determined previously [Costa et al. 1989]. The TTR genomic fragment used in this assay also contains sites for HNF-1, C/EBP, and HNF-4, none of which were protected by either the purified protein or the in vitro translation product, emphasizing that HNF-3A binds to a distinct site. All of these data confirm that the H3 clone encodes the sequence-specific DNA-binding protein HNF-3A, which we have purified.

The portion of the HNF-3A protein that is important for DNA binding was tested by transcription and translation of RNA that represented the full-length protein or subsections of the protein. Figure 7 shows the analysis of proteins made from a series of 3' deletions of the H3 clone. Figure 7A shows the products of in vitro translation after electrophoresis through an SDS-polyacrylamide gel. The total translation products were assayed for site-specific DNA-binding activity as shown in Figure 7B. Removal of up to 178 amino acids from the carboxy-terminal end did not affect site-specific DNA-binding activity [PvuII and deletion 7; a nonspecific band between arrows 1 and 2 arises from reticulocyte proteins]. Deletion of an additional 38 amino acids effectively abolishes the specific binding activity [deletion 14]. These experiments plus 5' deletions that were translated in vitro (data not shown) limit the DNA-binding region to amino acids 124–288 and show that amino acids 251–288 contain an essential part of the binding domain (Fig. 5, bottom).

Comparisons of the HNF-3A sequence with those proteins in the Genbank data base revealed no homology to any known DNA-binding proteins. Thus, HNF-3A contains a novel DNA-binding structure that remains to be characterized fully. A distant but interesting homology was noted with the Drosophila Notch protein. The sequence of Notch predicts a large transmembrane protein with multiple epidermal growth factor (EGF)-like repeats in the extracellular domain. The region of homology with HNF-3A extends from nucleotide residues 2230 to 2600, which encodes the intracellular domain of the protein. Recently, two yeast regulatory proteins, SW114 and SW116, have been found to exhibit distant homology to Notch but in yet another region of this protein’s intracellular domain. At this point, it is only speculation to suggest any related effector function between any of these evolutionarily distant proteins.

Limited distribution of HNF-3A is based on transcriptional control of the gene

Because the presence of HNF-3A-binding activity was not detected in brain, spleen, or kidney extracts but was restricted to liver extracts, we next sought to establish the level of such regulation. The steady-state levels of HNF-3A mRNA in several rat tissues were tested by Northern blot analysis. Two predominant mRNA species were found in liver with sizes of 2 and 2.2 kb. A smaller amount of the same mRNA was present in small intestine. These mRNAs were not detected in spleen, brain, and kidney RNA (Fig. 8A). We do not yet know the basis for the expression of the two different mRNA species.

Run-on transcription rates determined from isolated nuclei from liver, brain, and kidney showed that the transcription rate is at least 20-fold greater in liver than in either of the other tissues studied (Fig. 8B). These data show that the HNF-3A gene, itself, is expressed in a tissue-specific manner and that this expression is regulated at the level of transcription.

Discussion

The initial issues to settle in the study of hepatocyte-specific transcription were (1) the number of different regulatory proteins involved and (2) whether any of these proteins were found only (or mainly) in hepatocytes. We now have at least some preliminary answers to these questions. The regulatory regions of several genes have been explored thoroughly, and each shows multiple sites of DNA–protein interaction required for hepatocyte-specific expression; several of these proteins give evidence of being present mainly in the liver of
Figure 5. Nucleic acid sequence of H3 cDNA encoding HNF-3A. The predicted amino acid sequence of the protein is shown below the nucleotide sequence beginning with the initiator AUG at base 58. The two peptide sequences matching that obtained by direct protein sequencing of purified HNF-3A peptides are underlined. Below the sequence is a diagram of the residues found by deletion analysis to be important for DNA binding (see Fig. 7).
We have found that the HNF-3 site in TTR gene is indeed occupied specifically in liver nuclei. Further, footprint method (J. Mirkovitch and J.E. Darnell, in prep., 1990), there is no resemblance to well-recognized HNF-3 site. In separate experiments, using a new in vivo abundant liver protein that recognizes the characterized oligonucleotide site, the HNF-3A protein is the most efficient for activation of the globin promoter in transfected adult animals (for review, see Johnson 1990). Two of these proteins have been cloned, and each is part of a multigene family. HNF-1 [LF-B1] is a homeo domain protein (Frain et al. 1989; Baumhueter et al. 1990; Finney 1990), and C/EBP is the original member of the leucine zipper family of proteins (Landschulz et al. 1988). This report adds a third cloned cDNA encoding a protein found mainly or only in the liver and interacting with positive cis-acting regulatory regions of several genes expressed in hepatocytes. The newly cloned gene HNF-3A encodes a novel protein whose sequence is clearly not related to either HNF-1 or C/EBP. Even within the ~150 amino acids that direct DNA binding (124–288), there is no resemblance to well-recognized DNA-binding motifs. There are many proline and glycine residues that would limit α-helix formation to relatively short regions; thus, HNF-3A is a candidate for a new class of site-specific DNA-binding proteins.

As tested by gel mobility-shift analysis with a specific oligonucleotide site, the HNF-3A protein is the most abundant liver protein that recognizes the characterized HNF-3 site. In separate experiments, using a new in vivo footprint method [J. Mirkovitch and J.E. Darnell, in prep., we have found that the HNF-3 site in TTR gene is indeed occupied specifically in liver nuclei. Furthermore, this site was shown earlier to be critical and sufficient for activation of the globin promoter in transfected HepG2 cells. However, the role of HNF-3A as a positive-acting protein has not been proved. Additional work will be required to determine whether HNF-3A can activate transcription or whether the other HNF-3-binding activities, HNF-3B or HNF-3C, represent distinct proteins with positive-acting function. In this regard we note that two mRNA bands are observed in Northern blot analysis, and recently, we have detected a clone of a related transcription or whether the other HNF-3-binding activity, HNF-3A (E. Lai, V.R. Prezioso, and J.E. Darnell, unpubl.). Whatever the outcome of detailed mechanisms of transcriptional regulation involving the HNF-3-binding site, the HNF-3A protein represents a site-specific DNA-binding protein whose cell distribution in animals is highly specific. Northern blot analysis failed to detect the mRNA for this protein in any tissue other than liver, and this limited distribution is based on transcriptional regulation. This latter result is most important because a major goal of this work is to discover how, in the course of development, liver-specific transcription is initiated and maintained. We showed previously that the gene encoding C/EBP was transcriptionally controlled (Xanthopoulos et al. 1989) and have recently also shown that HNF-4 is transcriptionally controlled (K. Xanthopoulos et al., in prep.). Thus, the present findings with the HNF-3A cDNA provide strong evidence that a bat-
Figure 7. DNA-binding domain is disrupted by a 38-amino-acid deletion between residues 253 and 291. (A) Analysis of in vitro translation products from 3' deletion constructs of HNF-3A by SDS-PAGE. The arrows indicate the positions of the major product translated in the reticulocyte lysate. (B) Binding of these products to the HNF-3-binding site by the gel mobility-shift assay. The arrows indicate the position of the protein-DNA complex observed with the largest three protein products. Deletions of the 3' end of the H3 clone were generated by restriction cleavage or by exonuclease III/mung bean nuclease digestion by use of the kit from Stratagene. The positions of the 3'-end base for each deletion are as follows: PvuII (1194), deletion 7 (923), deletion 14 (807). In vitro translation and analysis of the products were performed as described in the legend to Fig. 2. Twice as much reticulocyte lysate was assayed for deletion 7 compared to the others to compensate for variation in translation efficiency in this preparation. The band between complexes 1 and 2 arises from reticulocyte proteins.

Figure 8. Cell-specific expression of HNF-3A. (A) Northern analysis of mRNA from several different tissues probed with a fragment of HNF-3A cDNA. (I) Intestine; (L) liver; (K) kidney; (B) brain; (S) spleen. (B) Run-on transcription rate from nuclei isolated from three tissues. Run-on transcription rates were determined in isolated nuclei from rat tissues, as described (Xanthopoulos et al. 1989).

Materials and methods

Purification of HNF-3A

Rat liver nuclei were isolated by pelleting through a sucrose cushion (Gorski et al. 1986) and extracted with buffer containing 0.4 M KCl. After removal of the chromatin and other insoluble material by centrifugation, the supernatant was collected and frozen in liquid N₂. Extract from 32 rat livers (~60–70 ml) was thawed and diluted with HGDE buffer (20 mM HEPES [pH 8.0], 10% glycerol, 1 mM dithiothreitol, 0.1 mM EDTA) to adjust the conductivity of the extract to 300 μS. All subsequent steps were performed at 4°C. The diluted sample was held for 20 min and centrifuged at 10,000g for 20 min to remove any precipitate. The supernatant was applied to a 60-ml heparin–agarose column (Sigma). Protein was eluted with a linear NaCl gradient (0.2–0.8 M). Fractions producing the slowest migrating complex (complex A, Fig. 1) were pooled, and the protein was concentrated by precipitation with 50% ammonium sulfate. This sample was applied to a 300-ml Sepharcl-S300 gel filtration column (Pharmacia), equilibrated, and eluted with HGDE/0.2 M KCl. Active fractions were pooled, applied to a Mono S FPLC column (0.5 x 5 cm), and eluted with a salt gradient of KCl. Fractions containing between 0.3 and 0.4 M KCl were adjusted to 0.15 M KCl and applied to a 2-ml DNA-cellulose column. The binding activity was eluted (0.25–0.4 M KCl), diluted to 0.15 M KCl, and applied to 1 ml affinity resin containing multimers of the double-stranded binding site for HNF-3 coupled to Sepharose (Radonaga and Tjian 1986). The binding site oligonucleotide used was

TCGAGTTGACTAAGTCAATAATCAGAATCAGCAACTGATTCAGTTATTAGTCTTAGTCAGCT

Active fractions were eluted between 0.25 and 0.4 M NaCl. This material was passed over the affinity column two additional times. Final purification for amino-terminal sequencing was...
achieved by applying the sample to an SDS–polyacrylamide gel and electrophoresis to PVDF membrane [Immobilon, Millipore]. The membrane was stained with Coomassie Blue, and the protein band of 50 kD was excised.

**Purification of HNF-3C**

The purification was essentially as described for HNF-3A, except that the activity monitored was gel mobility-shift complex C (Fig. 1). Fractions between 0.25 and 0.3 M KCl were collected on Immobilon. Purification was essentially as described for HNF-3A, except that the activity monitored was gel mobility-shift complex C (Fig. 1). Fractions between 0.25 and 0.3 M KCl were collected on Immobilon. The high-stringency wash was performed with probe 144 on a 5% native acrylamide gel, and then the entire gel was exposed to DNA cleavage by Cu-phenanthroline, as described by Kuwabara and Sigman (1987). The free and bound probe was located by autoradiography of the treated gel and electroeluted onto DEAE-cellulose. The cleaved probe was eluted from the membrane and analyzed on a 8% denaturing gel. The dried gel was exposed to X-ray film at −80°C with an intensifying screen.

**RNA analysis**

RNA was isolated from rat tissues by the method of Chirgwin et al. (1979). Poly[A]+ RNA was selected by chromatography on oligo(dT) cellulose. Five micrograms of poly[A]+ RNA from each tissue was separated on a glyoxal–DMSO (dimethyl sulfoxide) 1% agarose gel and transferred to nitrocellulose [Schleicher & Schuell] by electrophoresis. Detection of the HNF-3A mRNA was achieved with a probe containing bases 1–910 of the H3 clone in 50% formamide, 5 × SSPE, 5 × Denhardt’s solution, and 100 μg/ml salmon sperm DNA at 42°C. The high-stringency wash was performed in 0.1 × SSPE/0.1% SDS at 65°C for 10 min. The blot was exposed to X-ray film at −80°C with two intensifying screens for 30 hr. RNA loading was evaluated by hybridization of the blot with the rat glyceraldehyde 3-phosphate dehydrogenase probe.

**Acknowledgments**

We thank A. Cernik for guidance with HPLC, D. Atherton and the Protein Sequencing Facility for amino acid sequence, R. Doolittle (UCSD) for protein sequence comparisons, W. Chen, C.F. Kuo, K. Xanthopoulos, and M. Ehrlich for help in the RNA analysis, and R.E. Paulson and F. Sladek for helpful discussions. This work was funded by National Institutes of Health (NIH) grant CA16006-16A, NIH training grant (AI07233-12) to R.H.C., and an NIH Endocrine Research training grant (DK073113-01) to E.L.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

**Note added in proof**

Sequence data described in this paper have been submitted to the EMBL/GenBank Data Libraries.

**References**

Baumhueter, S., D.B. Mendel, P.B. Conley, C.J. Kuo, C. Turk, M.K. Graves, C.A. Edwards, G. Courtois, and G.R. Crabtree. 1990. HNF-1 shares three sequence motifs with the POU domain proteins and is identical to LF-B1 and APF. Genes Dev. 4: 372–379.

Cereghini, S., M. Raymondjean, A.G. Carranca, P. Herbomel, and M. Yaniv. 1987. Factors involved in control of tissue-specific expression of albumin gene. Cell 50: 627–638.

Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter.
Lai et al.

1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 24: 5294–5299.

Costa, R.H., D.R. Grayson, and J.E. Darnell, Jr. 1989. Multiple hepatocyte-specific nuclear factors function in the regulation of the transthyretin and α1-antitrypsin genes. *Mol. Cell. Biol.* 9: 1415–1425.

Costa, R.H., E. Lai, and J.E. Darnell, Jr. 1986. Transcriptional control of the mouse prealbumin (transthyretin) gene: Both promoter sequences and a distinct enhancer are cell specific. *Mol. Cell. Biol.* 6: 4697–4708.

Costa, R.H., E. Lai, D.R. Grayson, and J.E. Darnell, Jr. 1988. The cell-specific enhancer of the mouse transthyretin (prealbumin) gene binds a common factor at one site and a liver-specific factor(s) at two other sites. *Mol. Cell. Biol.* 8: 81–90.

Courtois, G., J.G. Morgau, L.A. Campbell, G. Fourel, and G.R. Crabtree. 1987. Interaction of a liver-specific nuclear factor with the fibrinogen and α1-antitrypsin promoters. *Science* 238: 688–692.

Derman, E., K. Krauter, L. Walling, C. Weinberger, M. Ray, and J.E. Darnell, Jr. 1986. Transcriptional control in the production of liver-specific mRNAs. *Cell* 23: 731–739.

DeSimone, V., G. Ciliberto, E. Hardon, G. Paonessa, F. Palla, L. Lundberg, and R. Cortese. 1988. Cis- and trans-acting elements responsible for cell-specific expression of the human α1-antitrypsin gene. *EMBO J.* 7: 2759–2766.

Finney, M. 1990. The homeodomain of the transcription factor LF-B1 has a 21 amino acid loop between helix 2 and helix 3. *Cell* 60: 5–6.

Frain, M., G. Swart, P. Monaci, A. Nicosia, S. Stampfli, R. Frank, and R. Cortese. 1989. The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. *Cell* 59: 145–157.

Gorski, K., M. Carneiro, and U. Schibler. 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. *Cell* 47: 767–776.

Herskowitz, I. 1989. A regulatory hierarchy for cell specialization in yeast. *Nature* 342: 749–757.

Johnson, P.F. 1990. Transcriptional activators in hepatocytes. *Cell Growth and Differ.* 1: 47–52.

Kadonaga, J. and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci.* 83: 5889–5893.

Kovesdi, I., R. Reichel, and J.R. Nevins. 1986. Identification of a cellular transcription factor involved in E1A trans-activation. *Cell* 45: 219–228.

Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283–292.

Kuwabara, M.D. and D.S. Sigman. 1987. Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: E. coli RNA polymerase-lac promoter complexes. *Biochemistry* 26(23): 7234–7238.

Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: A hypothetical structure common to a new class of DNA-binding proteins. *Science* 240: 1759–1764.

Lathe, R. 1985. Synthetic oligonucleotide probes deduced from amino acid sequence data. *J. Mol. Biol.* 183: 1–12.

Lichtsteiner, S., I. Wuarin, and U. Schibler. 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. *Cell* 51: 963–973.

Melton, D.A., P.A. Krug, M.R. Rebagliatti, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12: 7035–7056.

Monaci, P., A. Nicosia, and R. Cortese. 1988. Two different liver-specific factors stimulate in vitro transcription from the human α1-antitrypsin promoter. *EMBO J.* 7: 2075–2087.

Nasmyth, K. and D. Shore. 1987. Transcriptional regulation in the yeast life cycle. *Science* 237: 1162–1170.

Powell, D.J., J.M. Friedman, A.J. Oulette, K.S. Krauter, and J.E. Darnell, Jr. 1984. Transcriptional and post-transcriptional control of specific messenger RNAs in adult and embryonic liver. *J. Mol. Biol.* 179: 21–35.

Reue, K., T. Leff, and J.L. Breslow. 1988. Human apolipoprotein CIII is regulated by positive and negative cis-activating elements and tissue-specific protein factors. *J. Biol. Chem.* 263: 6857–6864.

Takahashi, Y., K. Kato, Y. Hayashizaki, T. Wakabayashi, E. Ohotsuka, S. Matsuki, M. Ikehara, and K. Matsubara. 1985. Molecular cloning of the human cholecystokinin gene by use of a synthetic probe containing deoxyinosine. *Proc. Natl. Acad. Sci.* 82: 1931–1935.

Xanthopoulos, K.G., J. Mirkovich, T. Decker, C.F. Kuo, and J.E. Darnell, Jr. 1989. Cell-specific transcriptional control of the mouse DNA-binding protein C/EBP. *Proc. Natl. Acad. Sci.* 86: 4117–4121.
HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally.

E Lai, V R Prezioso, E Smith, et al.

*Genes Dev.* 1990, 4:

Access the most recent version at doi:10.1101/gad.4.8.1427