OC-2, a Novel Mammalian Member of the ONECUT Class of Homeodomain Transcription Factors Whose Function in Liver Partially Overlaps with That of Hepatocyte Nuclear Factor-6*

(Received for publication, October 21, 1998)

Patrick Jacquemin‡, Vincent J. Lannoy‡§, Guy G. Rousseau, and Frédéric P. Lemaigre¶

From the Hormone and Metabolic Research Unit, Louvain University Medical School and the Christian de Duve Institute of Cellular Pathology, Avenue Hippocrate 75, B-1200 Brussels, Belgium

Transcription factors of the ONECUT class, whose prototype is hepatocyte nuclear factor (HNF)-6, are characterized by the presence of a single cut domain and by a peculiar homeodomain (Lannoy, V. J., Bürglin, T. R., Rousseau, G. G., and Lemaigre, F. P. (1998) J. Biol. Chem. 273, 13552–13562). We report here the identification and characterization of human OC-2, the second mammalian member of this class. The OC-2 gene is located on human chromosome 18. The distribution of OC-2 mRNA in humans is tissue-restricted, the strongest expression being detected in the liver and skin. The amino acid sequence of OC-2 contains several regions of high similarity to HNF-6. The recognition properties of OC-2 for binding sites present in regulatory regions of liver-expressed genes differ from, but overlap with, those of HNF-6. Like HNF-6, OC-2 stimulates transcription of the hnf-3β gene in transient transfection experiments, suggesting that OC-2 participates in the network of transcription factors required for liver differentiation and metabolism.

The identification of transcription factors has provided insight into the mechanisms of the control of gene expression. Some of these factors contain a DNA-binding region called the homeodomain. The homeodomain proteins are evolutionarily conserved and play an important role in cell differentiation and in morphogenesis (1–4). Several homeodomain proteins contain a second type of DNA-binding domain. This is the case for the proteins of the CUT superclass, in which the second DNA-binding domain is called cut because it was initially described in the Drosophila CUT protein (5–7). The CUT superclass comprises three classes (8). The CUX class, whose members have three cut domains, includes the Drosophila CUT protein and its mammalian homologs, namely human CDP, rat CDP-2, dog CLOX, and mouse CUX and CUX-2. The SATB class, whose members have two cut domains, includes the human homeodomain proteins called matrix attachment region-binding proteins or SATB (special AT-rich binding) proteins. A third class, called ONECUT because its members have a single cut domain, was identified (8) thanks to the cloning of rat hepatocyte nuclear factor-6 (HNF-6)† (9). The ONECUT class includes mammalian HNF-6 and four Caenorhabditis elegans cDNAs or open reading frames (ORFs) (8).

The proteins of the ONECUT class are characterized not only by their single cut domain, but also by a homeodomain with a peculiar amino acid composition. Homeodomains, which are 60 residues long, are organized in three α-helices (for a review on homeodomain-DNA interactions, see Refs. 10–12). The third helix, called the recognition helix, contacts the DNA and is crucial for sequence-specific binding. Within this helix, residue 48 is part of a hydrophobic core. Whereas residue 48 is a tryptophan in all known homeodomains, it is a phenylalanine in the ONECUT proteins. Residue 50 is also located in the recognition helix. Mutations at this position often lead to changes in the sequence specificity of DNA binding (13–15). This is consistent with the crystallization data that demonstrate that the amino acid at position 50 is in contact with bases (for a review, see Ref. 16). In the ONECUT proteins, a methionine is found at position 50. This amino acid is never found at this position in other homeodomains.

Our previous experiments showed that HNF-6 can bind to a number of DNA sites, which differ slightly in terms of sequence, and that the cut domain is required for binding of HNF-6 to all the sites tested (8). These experiments ascribed a dual role to the peculiar homeodomain of HNF-6. They showed that this domain is involved in DNA binding, but only for a subset of the sites recognized by HNF-6. They also showed that the homeodomain is involved in transcriptional activation, but only of those genes for which the binding of HNF-6 does not require its homeodomain. By mutational analysis, we demonstrated that phenylalanine 48 and methionine 50 play a role in this transcriptional activating function of the HNF-6 homeodomain (8). This work indicated that the linker region between the cut domain and homeodomain of HNF-6 is important for DNA binding (8). We identified two rat isoforms (HNF-6α and HNF-6β) that originate from the same gene by alternative splicing (17). HNF-6β (491 residues) is identical to HNF-6α (465 residues), except that it contains an insert of 26 amino acids in the linker region. These two isoforms differ in DNA

* This work was supported by grants from the Belgian State Program on Interuniversity Poles of Attraction, Prime Minister’s Office, Federal Office for Scientific, Technical, and Cultural Affairs; from the Délégation Générale Higher Education and Scientific Research of the French Community of Belgium; from the Fund for Scientific Medical Research (Belgium); and from the National Fund for Scientific Research (Belgium). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) Y18198.

‡ These two authors equally contributed to this work.

§ Recipient of a fellowship from the Fonds pour la Formation à la Recherche dans l’Industrie et l’Agriculture (Belgium).

¶ Senior Research Associate from the National Fund for Scientific Research (Belgium). To whom correspondence should be addressed.

Tel.: 32-2-764-7583; Fax: 32-2-762-7455; E-mail: lemaigre@horm.ucl.ac.be.

† This abbreviation was used: HNF, hepatocyte nuclear factor; ORF, open reading frame; RT-PCR, reverse transcription-polymerase chain reaction; TTR, transthyretin; PEPCk, phosphoenolpyruvate carboxykinase; GRU, glucocorticoid-responsive unit; bp, base pair(s); CORD, cone-rod dystrophy.

The abbreviations used are: HNF, hepatocyte nuclear factor; ORF, open reading frame; RT-PCR, reverse transcription-polymerase chain reaction; TTR, transthyretin; PEPCk, phosphoenolpyruvate carboxykinase; GRU, glucocorticoid-responsive unit; bp, base pair(s); CORD, cone-rod dystrophy.
As mentioned above, the ONECUT class contains several *C. elegans* members, but only one mammalian member, namely HNF-6. The DNA-binding domains of these *C. elegans* proteins display DNA binding properties similar to those of HNF-6 (8). This indicates that these properties have been evolutionarily conserved and are therefore important for basic regulatory processes that are common to nematodes and mammals. The DNA-binding domains of these *C. elegans* proteins display DNA binding properties similar to those of HNF-6 (8). This indicates that these properties have been evolutionarily conserved and are therefore important for basic regulatory processes that are common to nematodes and mammals. The DNA-binding domains of these *C. elegans* proteins display DNA binding properties similar to those of HNF-6 (8). This indicates that these properties have been evolutionarily conserved and are therefore important for basic regulatory processes that are common to nematodes and mammals.

**FIG. 1.** OC-2 is a novel mammalian member of the ONECUT class of homeodomain proteins. A, nucleotide sequence of OC-2 cDNA and deduced amino acid sequence. The cut domain and homeodomain are boxed. B, alignment of the amino acid sequences of human (h) OC-2, human HNF-6, *C. elegans* (ce) ORF R07D10.x, and human ORFs found in cosmid F21967 and fosmid F37502. Residues identical to those in OC-2 are on a black background; similar residues are indicated by gray shading. Boxed sequences correspond to regions whose structures predicted by the Match-Box algorithm are similar among all the sequences listed. Asterisks point to Phe-48 and Met-50 of the homeodomain.
existence of several members of the ONECUT class in C. elegans prompted us to search for mammalian members of this class that are distinct from HNF-6. We describe here a new member, which we call OC-2, of the ONECUT class.

**EXPERIMENTAL PROCEDURES**

**Oligonucleotides**—The polymerase chain reaction (PCR) primers were as follows: bACT5, 5'-GGCATCGTGATGGACTCCG-3'; bACT3, 5'-GCTGGAAGGTGGACAGCGA-3'; FM2.1, 5'-ATTTCCCAGCAGCTGGG-3'; FM2.2, 5'-TATCTGTTGTCCTCATTTGG-3'; FM2.2A, 5'-CAGCTAGGAATCCGGTGTC-3'; FM2.9, 5'-TGGGGCTGGCCAGCAGCT-3'; FM2.10, 5'-AAAGTCTGGGCACTTTGC-3'; HNF6.6, 5'-GCTAGACGAGCAAAATCACACTCC-3'; HNF6.5, 5'-AAACCAAAGACTTAGCTCACCTGCA-3'; VLIIA, 5'-CCCAAGCTTACCATGGCAATCAATACCAAGAGGTGGCTC-3'; and VLIIB, 5'-CGGGATCCTCACAGCCACTTCACATCCTCCG-3'. The primers used as probes in Southern blotting experiments were as follows: HOM2, 5'-CCYMGRYTSGYSTTYAC-3'; HOM1, 5'-CKNSYRTTCATRAARAA-3'; HNF6.3, 5'-AGGAAAGAGCAAGAACATGGGAAGG-3'; and HNF6.4, 5'-CTGCTCATCATTTGTCTTGCCGCAAG-3'. The double-stranded oligonucleotide probes used in electrophoretic mobility shift assays were as follows: HNF-4, 5'-AGGATAGATGATGATCTGGGA-3' (2394 to 2371 of the mouse hnf-4 promoter); HNF-3β, 5'-AGCTTAAGGCCCGATATTGATTTTTTTCTCC-3' (2150 to 2118 of the rat hnf-3β promoter); TTR, 5'-GTCTGAAAGTCAATAATCAGAAT-3' (2110 to 2087 of the mouse transthyretin promoter); PEPCK, 5'-CAAAGTTTAGTCAATCAAACGTTG-3' (2263 to 2240 of the rat phosphoenolpyruvate carboxykinase (pepck) gene promoter); and GRU, 5'-AAAAAAATCCATAACTTTCA-3' (sequence in the GRU of the first intron of rat 6-phosphofructo-2-kinase (pfk-2) gene A).

**Cloning of OC-2 cDNA**—A 193-base pair (bp)-long OC-2 probe was synthesized using the PCR primers FM2.1 and FM2.2 and the IMAGE clone 566080 as template. This clone contains an expressed sequence tag (GenBank™ accession number AA121823) derived from a human fetal retina cDNA library. The OC-2 probe was used to screen a HeLa cell cDNA library (kindly provided by J.-M. Garnier) prepared in λZap II (Stratagene) by hybridization at 42 °C in 6×SSC and 50% formamide. Filters were washed at 55 °C in 3×SSC, and positive clones were isolated according to the supplier's instructions. Clones lacking the 5'-end of the cDNA were obtained. To find this 5'-end, a human genomic λEMBL3 library was screened with a probe containing the OC-2 cut domain. This probe was obtained by PCR using primers VLIIA and VLIIB and a HeLa cell OC-2 cDNA clone as template. A 4.5-kilobase pair-long EcoRI-EcoRI fragment was obtained from a purified genomic clone and subcloned in pBluescript KS1 (Stratagene). A cDNA containing a full-length coding sequence was constructed by fusing 585 bp of the genomic subclone (nucleotides 1–585 in Fig. 1A) to 1069 bp of a HeLa cell OC-2 cDNA clone (nucleotides 586–1654 in Fig. 1A). The genomic portion used to construct the fully coding cDNA is expressed as mRNA as confirmed by RT-PCR with liver or skin RNA. These RT-PCR experiments were performed with the Titan RT-PCR system (Boehringer Mannheim) in the presence of 2M betaine using primers FM2.9 and FM2.10. The sequence of the two DNA strands was determined by automated DNA sequencing using the dideoxy chain termination method. DNA and protein sequence analysis was performed using the GCG software package (Genetics Computer Group, Inc., University of

**FIG. 1—continued**
Characterization of the ONECUT Transcription Factor OC-2

To identify predicted structurally conserved regions, multiple sequence alignments were performed using the Match-Box algorithm with matrix blosum62 (18, 19).\(^2\)

**Radiation Hybrid Mapping**—The chromosomal location of the human OC-2 gene was characterized by radiation hybrid mapping using the GeneBridge 4 Radiation Hybrid panel (Research Genetics). This panel of 93 human × hamster somatic cell hybrids was screened by PCR (35 cycles at 94 °C for 1 min, 57 °C for 1 min, and 72 °C for 1 min) with primers FM2.1 and FM2.2A in 50-μl incubations and visualized on an ethidium bromide-stained agarose gel. The data were analyzed using the Whitehead Institute/Massachusetts Institute of Technology Center for Genome Research radiation hybrid map.\(^3\)

**RT-PCR**—To determine the tissue distribution of OC-2 mRNA, RT-PCR was performed (GeneAmp kit, Perkin-Elmer) with 500 ng of total RNA isolated from human tissues. The integrity of the RNA preparations was tested by amplification of a 610-bp-long β-actin CDNA fragment with primers bACT5 and bACT3. cDNA (193-bp-long fragment) was amplified with primers FM2.1 and FM2.2. HNF-6 cDNA (387-bp-long fragment) was amplified with primers HNF6.5 and HNF6.6. Primers FM2.2 and HNF6.6 were used in the reverse transcription step. The specificity of the amplified products was verified in Southern blotting experiments as described (20) using, as \(^3^2^P\)-labeled probes, PCR fragments obtained with primers HOM1 and HOM2 for the identification of OC-2 products and with primers HNF6.3 and HNF6.4 for the identification of HNF-6 products.

**Expression Vectors and Reporter Constructs**—The expression vector pCMV-HNF6a contains an AccI-EcoRI fragment of the HNF-6a CDNA, derived from pSP-HNF6a (9), cloned downstream of a cytomegalovirus promoter. pECE-HNF6a has been described (9). pXJ42-OC2 contains a fully coding human OC-2 cDNA cloned in the Xhol and EcoRI sites of pXJ42 (21). The pHNF6/HNF3\(\beta\)mut-luc, pHNF3\(\beta\)-luc, pHNF3\(\beta\)-luc, and pHNF4/0.7-luc reporter vectors have been described (8, 22, 23). The internal control vector pRLnull contains the ptk-2 promoter (−138 to +86) cloned in pRLnull (Promega).

**Transfections and Cell Extracts**—Rat hepatoma FTO-2B cells were grown in Dulbecco’s modified Eagle’s medium/ham’s F-12 medium supplemented with 10% fetal calf serum. Cells (1 × 10⁵ cells/6-cm dish) were transfected in Dulbecco’s modified Eagle’s medium without fetal calf serum by lipofection using LipofectAMINE PLUS (Life Technologies, Inc.), 400 ng of reporter construct (pHNF6/HNF3\(\beta\)mut-luc, pHNF3\(\beta\)-luc, pHNF3\(\beta\)-luc, or pHNF4/0.7-luc) were incubated on ice for 20 min in a final volume of 20 μl containing 10 μg HEPES (pH 7.6), 1 mM dithiothreitol, 1 mM MgCl₂, 50 mM KCl, 10% (v/v) glycerol, 4 μg of poly(dI-dC), and the \(^3^2^P\)-labeled probe (30,000 cpm). The samples were loaded on a 7% acrylamide gel (29:1 acrylamide/bisacrylamide ratio) in 0.25× Tris borate/EDTA buffer and electrophoresed at 200 V.

**Electrophoretic Mobility Shift Assays**—COS-7 cell lysates (5 μl) were incubated on ice for 20 min in a final volume of 20 μl containing 10 μg HEPES (pH 7.6), 1 mM dithiothreitol, 1 mM MgCl₂, 0.5 mM EGTA, 50 mM KCl, 10% (v/v) glycerol, 4 μg of poly(dI-dC), and the \(^3^2^P\)-labeled probe (30,000 cpm). The samples were loaded on a 7% acrylamide gel (29:1 acrylamide/bisacrylamide ratio) in 0.25× Tris borate/EDTA buffer and electrophoresed at 200 V.

![Graph showing tissue distribution of OC-2 and HNF-6](image)

**Fig. 3. Tissue distribution of human OC-2 and human HNF-6.** The tissue distribution of OC-2 and HNF-6 was determined by RT-PCR, and amplified products were identified by Southern blotting. As a control for RNA integrity, β-actin mRNA was amplified by RT-PCR, and the products were visualized on a nondenaturing acrylamide gel stained with ethidium bromide. Ki, kidney; Lu, lung; Mu, skeletal muscle; Br, brain; Sp, spleen; Te, testis; He, heart; My, myometrium; Li, Liver; Bl, urinary bladder; Sk, skin; PBL, peripheral blood lymphocytes.

\(^2\) http://www.fundp.ac.be/sciences/biologie/bms/matchbox_submit.html.

\(^3\) http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl.
RESULTS AND DISCUSSION

Identification of OC-2—During a search for HNF-6-related proteins in GenBank™ using the Basic Local Alignment Search Tool (BLAST) (24), we found an expressed sequence tag (GenBank™ accession number AA121823) from a fetal human retina cDNA library that contains a partial cDNA showing significant similarity to the HNF-6 homeobox. Two primers (FM2.1 and FM2.2; see “Experimental Procedures”) whose sequences were based on the expressed sequence tag were used to screen, by PCR, two cDNA libraries, one from adult human retina and one from HeLa cells. Identical 193-bp-long bands were seen with the two libraries, yielding a probe that was used to screen the HeLa cell cDNA library by phage plaque hybridization. Six independent clones were obtained. All these clones lacked the 5'-end of the coding sequence, but contained nearly identical 5'-extremities. This probably reflects the presence of secondary structures in the mRNA that blocked reverse transcription during the preparation of the library. Rapid amplification of cDNA ends/PCR failed to amplify the missing portion, which turned out to be GC-rich (see below and Fig. 1A). To clone the 5'-end, we therefore screened a human genomic DNA library with a probe encompassing nucleotides 1027–1214 (Fig. 1A). An additional 5'-sequence was obtained (nucleotides 1–585 in Fig. 1A), and RT-PCR with high processivity polymerase on total RNA from human liver and skin confirmed that this sequence is expressed as mRNA. The fully coding cDNA and the derived 485-amino acid-long sequence are shown in Fig. 1A. The ATG codon at nucleotides 88–90 lies within a perfect Kozak consensus sequence, strongly suggesting that it codes for the initiator methionine. Sequence comparisons showed the presence of a single cut domain (residues 314–379) and a homeodomain (residues 407–466). The homeodomain was typical of the ONECUT class, with a phenylalanine at position 48 and a methionine at position 50. We therefore named the novel protein OC-2 (for ONECUT-2).

Fig. 1B shows an alignment of the amino acid sequences of OC-2 and HNF-6. This alignment shows that the sequences of the cut domain and homeodomain of OC-2 are 97 and 87% identical to those of HNF-6, respectively, and that HNF-6 and OC-2 display 92% amino acid sequence identity over the 157 residues (boxed in Fig. 1B) that encompass the entire DNA-binding region. There are additional regions of similarity between OC-2 and HNF-6 (Fig. 1B), such as a serine-rich tail of 19 residues in OC-2 and of 21 residues in HNF-6. The linker between the cut domain and homeodomain has the same length (27 residues) in OC-2 and HNF-6, and the amino acid sequence of this linker is quasi-identical (83%) in both proteins. A search by RT-PCR with RNA from tissues expressing OC-2 (see below) and primers located in the cut domain and homeodomain failed to provide evidence for an OC-2 isoform containing an alternatively spliced insert between these domains, as is the case for HNF-6β (data not shown). As in HNF-6, a polyhistidine
Characterization of the ONECUT Transcription Factor OC-2

The underlined sequences correspond to the HNF-6-binding site consensus sequence (8).

| Target gene | Site sequence | Transcription factor binding |
|-------------|---------------|-----------------------------|
| hnf-3B     | -112GAAAAAAAATACATCGGAGCT-144 | + + + + |
| hnf-4      | -395AGCTAGAATGCTAGATCCTTGGAG-371 | - + + + |
| Transthyretin | -11GCTCTGCTTAGCTAAATCCAGAAAT-97 | + + + |
| pepck      | -26CAGAGTTTGATGTTCCACAGGTG-240 | + + + + |
| p/κ-2 GRU  | (1st intron) AACAAAAAAATCCATATGTTCA | + + + + |

4 V. J. Lannoy, G. G. Rousseau, and F. P. Lemaigre, manuscript in preparation.

5 A. M. Müller, J. Ek, S. M. Duriaux, S. A. Urhammer, J. O. Clausen, H. Eiberg, T. Hansen, G. G. Rousseau, F. P. Lemaigre, and O. Pedersen, submitted for publication.

6 C. Pierreux, J. Stafford, D. Demonte, D. K. Scott, J. Vandenhaute, R. M. O’Brien, D. K. Granner, G. G. Rousseau, and F. P. Lemaigre, submitted for publication.

tract occurs in OC-2 at about the same distance upstream of the cut domain. A sequence of 14 consecutive glycines (residues 22–35) is present in OC-2. Such stretches of identical amino acids have been found in many homeodomain proteins (25) and may be functionally important. Indeed, alterations of a polyalanine stretch in the HOXD13 protein causes polydactyly (26). Yet another conserved region (92% identity between OC-2 and HNF-6) is a serine/threonine- and proline-rich sequence (STP box) of 24 residues located just upstream of the polyhistidine tract. Our experiments on HNF-6α suggest that this STP box plays a role in transcriptional activation.4 Fig. 1B also shows an alignment of the amino acid sequences of OC-2 and HNF-6α with that of the C. elegans ONECUT gene product closest to these mammalian proteins, namely R07D10.x (an ORF found on cosmid C17H12). One sees that not only the homeodomain and the cut domain, but also the STP box, have been conserved in the ONECUT class since the nematodes.

Recent data bank searches showed that the human genome contains sequences that code for additional putative ONECUT proteins very similar to HNF-6 and OC-2. The amino acid sequences of two ORFs, one present in fosmid F37502 (GenBank accession number AC004755) and one present in cosmid F21967 (GenBank accession number AC005256), are shown in Fig. 1B. Fosmid F37502 and cosmid F21967 contain sequences present on the same chromosome (human chromosome 19) and separated by a 1-kilobase pair gap. The distance between the two ORFs shown in Fig. 1B is 19 kilobase pairs. The ORF of fosmid F37502 contains a typical ONECUT homeodomain with a phenylalanine at position 48 and a methionine at position 50. The ORF of cosmid F21967 contains a cut domain that is quasi-identical to those of the bona fide ONECUT proteins. It also contains an STP box and a polyglycine stretch. It could be that cosmid F21967 and cosmid F37502 each contain one-half of a single ONECUT gene whose organization resembles that of the hnf-6 gene, namely with a very long intronic region between the cut domain and homeodomain (17). In this case, the two corresponding ORFs shown in Fig. 1B would belong to two parts of the same OC-3 protein. If these ORFs do not correspond to pseudogenes and if these ORFs are expressed as peptides, these peptides very likely correspond to ONECUT protein(s) different from HNF-6 and OC-2.

Chromosomal Localization of OC-2—The chromosomal localization of OC-2 in humans was determined by radiation hybrid mapping. PCR amplifications were performed, and the data were analyzed as indicated under “Experimental Procedures.” This showed that the OC-2 gene is located on chromosome 18, 10.31 cR3000 telomeric from marker WI-8740 and 9.2 cR3000 centromeric from marker CHLC.GATA30B03 (lod score > 3.0). Human HNF-6 is located on chromosome 15q21.1-21.2 (27), and, as said above, the putative human OC-3 gene is located on chromosome 19. These data show that in humans, the genes coding for the homeodomain proteins of the ONECUT class are not located in a cluster.

Like OC-2, marker D18S69 is located between markers WI-8740 and CHLC.GATA30B03. Since D18S69 is cytogenetically located at 18q21.1-18q21.2 (28), we conclude that OC-2 is most probably also located cytogenetically at 18q21.1-18q21.2. Deletions encompassing 18q21.1-21.3 are associated with retinal cone-rod dystrophy (CORD)-1 (29). Considering that OC-2 is expressed in the retina (see “Cloning of OC-2 cDNA,” above) and that the loss of the homeodomain-containing transcription factor CRX causes the related syndrome CORD2 (30, 31), further work should identify the retinal cell types expressing OC-2 and the possible association between CORD1 and OC-2 dysfunctions.

OC-2 Gene Expression Is Tissue-restricted—To determine the tissue distribution of OC-2, we screened, by RT-PCR, 12 human tissues for expression of its mRNA. The tissue distribution of HNF-6 mRNA, not yet studied in humans, was analyzed in parallel. The PCR products were subjected to Southern blot analysis using overlapping radioactive PCR probes. As a control, β-actin mRNA was amplified from the same RNA preparations. The data (Fig. 2) show that OC-2 mRNA is abundant in the liver and skin. Lower amounts were found in the testis, brain (occipital cortex), and urinary bladder. The kidney yielded a very weak signal. The expression of HNF-6 mRNA was strongest in the liver and clearly detectable in the testis and skin. We conclude that OC-2 has a tissue-restricted pattern of expression that differs from, but overlaps with, that of HNF-6.

OC-2 and HNF-6 Have Distinct but Overlapping DNA Binding and Transcriptional Activation Properties—The remarkable amino acid sequence similarity of the two DNA-binding domains of OC-2 and HNF-6α and of the linker region between these domains suggested that the two proteins display similar DNA binding specificities. HNF-6 binds to two types of sequences. Binding to the first type, exemplified by a probe derived from the transthyretin gene (TTR probe), requires both the cut domain and homeodomain of HNF-6. Binding to the second type of sequence, exemplified by a probe derived from the hnf-3β gene (HNF-3β probe), requires only the cut domain of HNF-6 (8). We therefore tested, by electrophoretic mobility shift assay, whether OC-2 binds to these two probes. This was the case (Fig. 3, lanes 3 and 8 versus lanes 1 and 6). Binding of OC-2 was specific as shown in incubations containing an excess of unlabeled probe (Fig. 3, lanes 4 and 9). Cross-competitions suggested that OC-2 binds with higher affinity to the HNF-3β probe than to the TTR probe (Fig. 3, lanes 5 and 10). In contrast, HNF-6α bound equally well to the two probes, as shown in Fig. 3 (lanes 2 and 7) and by our earlier results (8).

HNF-6 regulates the transcription of genes involved in glucose metabolism (9, 32). As shown in Fig. 3 (lanes 14–19), we
found that OC-2 and HNF-6α bind to the same cis-acting sequences in the pepck (PEPCK probe) and pfk-2 (GRU probe) genes. Like HNF-6, OC-2 is therefore expected to play a role in the regulation of liver gluconeogenesis and glycolysis.

Experiments on differentiated hepatoma lines and on embryoid bodies provided evidence for a network of liver-enriched transcription factors that is required for hepatocyte differentiation and morphogenesis (33–38). HNF-3β is an essential component of this network. It stimulates expression of hnf-4. On the other hand, HNF-4 activates transcription of the hnf-1 gene. Our earlier experiments indicated that HNF-6 binds to, and controls the expression of, the hnf-3β and hnf-4 genes (22).

We therefore determined whether OC-2 binds to the same sites as HNF-6 in these two genes. While OC-2 bound to the same HNF-3β probe as HNF-6 (see above), OC-2 did not bind to the HNF-6-binding site derived from the hnf-4 gene promoter (HNF-4 probe) (Fig. 3, lanes 11–13). The discrimination of OC-2 between the HNF-3β and HNF-4 probes is similar to that described earlier for HNF-6β (8). However, OC-2 binds to a probe (PEPCK) to which HNF-6β binds very poorly. Therefore, OC-2 has a DNA binding specificity that differs from, but overlaps with, that of HNF-6α and HNF-6β (Table I).

To study if OC-2 may be involved in the network of liver transcription factors by controlling transcription of the hnf-3β gene, rat hepatoma FTO-2B cells were cotransfected with an OC-2 gene promoter, rat hepatoma FTO-2B cells were cotransfected with an OC-2 expression vector and a construct containing the luciferase reporter under the control of six copies of the HNF-6/OC-2 site found in the pepck gene promoter (pHNF3β-luc) or the same promoter in which the HNF-6/OC-2 site has been destroyed by mutation (pHNF3βmut-luc). Fig. 4B shows that overexpression of OC-2 or HNF-6α stimulated transcription from the wild-type promoter 2–3-fold, but was without effect on the mutated promoter. Consistent with earlier results (22), the mutated promoter displayed a 50% reduction in basal activity compared with the wild-type promoter, probably as a result of the loss of action of endogenous ONECUT transcription factors. Finally, as expected from the lack of binding of OC-2 to the probe derived from the hnf-4 gene (Fig. 3), overexpression of OC-2 did not activate the hnf-4 promoter (pHNF4-0.7-luc) (Fig. 4B) in transfection experiments, whereas HNF-6α did. We conclude from these transfection experiments that OC-2 stimulates transcription of the hnf-3β gene. Both HNF-6 and OC-2 might therefore be key players in the network of liver transcription factors by controlling a different set of genes.

Conclusions—Human OC-2 is a mammalian member of the ONECUT class of homeodomain transcription factors whose sequence is very similar to that of HNF-6. On the basis of their respective tissue distribution, DNA binding specificities, and transcriptional activation properties, HNF-6 and OC-2 have different but overlapping functions. The expression of OC-2 in the retina and its chromosomal localization raise the question of the involvement of OC-2 in CORD1. In liver, OC-2 is expected to participate in the network of transcription factors that regulates differentiation and morphogenesis.

Acknowledgments—We thank J.-M. Garnier for the HeLa cell cDNA library, J. Nathans for the retina cDNA library, I. Davidson for pX342, R. Costa for pHNF6/HNF3β(-)–TATA-luc, F. Brasseur for human RNA samples, the United Kingdom Human Genome Mapping Project Resource Center for the IMAGE clone 566080, and S. Durviaux and S. Neou for expert technical assistance.

REFERENCES

1. Kenyon, C. (1994) Cell 78, 175–180
2. Lawrence, P. A., and Morata, G. (1994) Cell 78, 181–189
3. Krumlauf, R. (1994) Cell 78, 191–201
4. Duboule, D. (ed) (1994) Guidebook to the Homeobox Genes, Oxford University Press, Oxford
5. Blochlinger, K., Bodmer, R., Jack, J., Jan, L. Y., and Jan, Y. N. (1988) Nature 333, 629–634
6. Harada, R., Dufort, D., Denis-Larose, C., and Nepveu, A. (1994) J. Biol. Chem. 269, 2062–2067
7. Anders, V., Chiara, M., and Madhavi, V. (1994) Genes Dev. 8, 245–257
8. Lannoy, V. J., Burglin, T. R., Rousseau, G. G., and Lemaigre, F. P. (1998) J. Biol. Chem. 273, 13552–13562
9. Lemaigre, F. P., Dubrux, S. M., Truong, O., Lannoy, V. J., Hsuun, J. A., and Rousseau, G. G. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 9460–9464
10. Kornberg, T. B. (1993) J. Biol. Chem. 268, 26813–26816
11. Gehring, W. W., Qian, Y. Q., Billiet, M., Furukado-Tokunaga, K., Schier, A. F., Rosendez-Perez, D., Affolter, M., Otting, G., and Wuthrich, K. (1994) Cell 78, 211–223
12. Wolberger, C. (1996) Curr. Opin. Struct. Biol. 6, 62–68
13. Hanes, S. D., and Brent, R. (1989) Cell 57, 1275–1283
14. Aresta, S. E., and Sauer, R. T. (1994) Biochemistry 33, 9187–9194
15. Stepheenko, A. G., Luchina, N. N., and Pankratova, E. V. (1997) Nucleic Acids Res., 25, 2847–2853
16. Treisman, J., Harris, E., Wilson, D., and Deplanck, C. (1992) Biosciences 14, 145–150
17. Rastegar, M., Spirir, C., Rousseau, G. G., and Lemaigre, F. P. (1998) Biochem. J. 334, 565–569
18. Depiereux, E., and Feytmans, E. (1992) Comput. Appl. Biosci. 8, 501–509
19. Depiereux, E., Baudoux, G., Briffeul, P., Regirister, I., De Bolle, X., Vinals, C., and Feytmans, E. (1997) Comput. Appl. Biosci. 13, 249–256
20. Yamanaka, S., and Nishizuka, S. (1997) J. Biol. Chem. 272, 12928–12937
21. Xia, J. H., Davidson, I., Matthes, H., Garnier, J.-M., and Chambon, P. (1991) Cell 65, 551–568
22. Landry, C., Clotman, F., Hiski, T., Oda, H., Picard, J. P., Lemaigre, F. P., and Rousseau, G. G. (1997) Dev. Biol. 15, 247–257
23. Rausa, F., Samadani, U., Ye, H., Lim, L., Fletcher, C. F., Jenkins, N. A., and Weinberg, R. A. (1997) Dev. Biol. 192, 228–236
24. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
25. Burglin, T. R. (1994) in Guidebook to the Homeobox Genes (Duboule D., ed) pp. 25–71, Oxford University Press, Oxford
26. Murakagi, Y., Mundlos, S., Upton, J., and Olsen, B. R. (1996) Science 272, 548–551
27. Vaisse, C., Kim, J., Espinoza, R. III., Le Beau, M. M., and Stoffel, M. (1997) Diabetes 46, 1364–1367
28. Papadimitrakopoulou, V. A., Oh, Y., El-Naggar, A., Izzo, J., Clayman, G., and Mao, L. (1998) Clin. Cancer Res. 4, 539–544
29. Wurzburg, M., Sjo, O., Traenbergh, L., and Fledelius, H. (1991) Am. J. Med. Genet. 39, 288–293
30. Furukawa, T., Morrow, E. M., and Cepko, C. L. (1997) Mol. Cell. Biol. 17, 543–553
31. Samadani, U., and Costa, R. H. (1996) Mol. Cell. Biol. 16, 6273–6284
32. Kuo, C. J., Conley, P. B., Chen, L., Sladek, F. M., Darnell, J. E., Jr., and Crabtree, G. R. (1992) Nature 355, 457–461
33. Bulla, G. A., and Fournier, R. E. (1994) Mol. Cell. Biol. 14, 7086–7094
34. Cereghini, S. (1996) FASEB J. 10, 267–282
35. Spaff, G. F., and Weiss, M. C. (1997) Mol. Cell. Biol. 17, 1913–1922
36. Spaff, G. F., and Weiss, M. C. (1998) J. Cell Biol. 140, 935–946
37. Duncan, S. A., Navas, A. M., Dufort, D., Rossant, J., and Stoffel, M. (1998) Science 281, 692–695