The HOXC11 Homeodomain Protein Interacts with the Lactase-Phlorizin Hydrolase Promoter and Stimulates HNF1α-dependent Transcription

(Received for publication, February 17, 1998)

Cathy Mitchelmore‡, Jesper T. Troelsen, Hans Sjöström, and Ove Norén

From the Department of Medical Biochemistry and Genetics, Biochemical Laboratory C, The Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

The lactase-phlorizin hydrolase (LPH) gene is expressed specifically in the enterocytes of the small intestine. LPH levels are high in newborn mammals, but decrease after weaning. We have previously suggested that the promoter element CE-LPH1, located at −40 to −54, plays an important role in this down-regulation, because the DNA binding activity of a nuclear factor that binds to this site is present specifically in small intestinal extracts and is down-regulated after weaning. In an effort to clone CE-LPH1-binding factors, a yeast one-hybrid genetic selection was used, resulting in the isolation of a partial cDNA encoding the human homeodomain protein HOXC11. The full-length HOXC11 sequence was obtained by rapid amplification of cDNA ends. It was shown in a yeast assay and by electrophoretic mobility shift assay that HOXC11 binds to the CE-LPH1 element with similar specificity to the endogenous intestinal factor. Two HOXC11 transcript sizes were identified by Northern blot analysis. The larger transcript (2.1 kilobase pairs) is likely to contain a translational start site in good context and is present in HeLa cells. The shorter 1.7-kilobase pair transcript, present in HeLa and Caco-2 cells, probably encodes a protein lacking 114 amino acids at the N-terminal end. Both forms of HOXC11 potentiate transcriptional activation of the LPH promoter by HNF1α. The expression of HOXC11 mRNA in human fetal intestine suggests a role in early intestinal development.

Lactase-phlorizin hydrolase (LPH) is a membrane-bound small intestinal enzyme that hydrolyzes the lactose present in milk. LPH is highly expressed in the newborn mammal, but is down-regulated after weaning. However, in some humans, mainly of Northern European descent, LPH activity remains high throughout adult life. In other humans, down-regulation of LPH leads to milk intolerance in adolescence and adult life, a clinical condition referred to as adult-type hypolactasia. Several studies have indicated that the down-regulation of LPH activity in mammals is mainly at a transcriptional level, although post-transcriptional mechanisms play a modifying role (1–4). A 1-kb promoter region of the pig LPH gene is able to drive appropriate tissue-specific and developmental regulation of a linked reporter gene, suggesting that the cis elements required for transcriptional regulation are located within this region (3).

Previous work in our group has sought to identify the cis elements of the lactase gene promoter, and the corresponding DNA-binding factors. A factor present in nuclear extracts prepared from pig small intestine was shown to bind to the CE-LPH1 site at −40 to −54 in the promoter, adjacent to the putative TATA-box. This DNA binding activity, termed NF-LPH1, was found to be intestine-specific and was present at higher levels in newborn pigs than in adult animals (5). NF-LPH1 is also present in Caco-2 cells (6), a cell line derived from a human colon carcinoma, which express LPH when allowed to differentiate (7). By using competition electrophoretic mobility shift assays (EMSA) and photoaffinity-labeling, it was suggested that NF-LPH1 in Caco-2 nuclear extracts is functionally related to a nuclear factor binding to the SIF1 element of the sucrase-isomaltase gene. Both factors recognize a TTTA(T/C) core sequence (6).

The SIF1-binding protein was recently cloned from mouse and identified as a homeodomain protein, Cdx-2 (8). The hamster homologue, termed Cdx-2 or shCdx-2, was previously isolated as a factor binding to an insulin gene promoter element (9). We have shown that Cdx-2 also binds to CE-LPH1 and activates LPH gene expression, dependent on a functional CE-LPH1 element, in co-transfection studies (10). The ability of Cdx-2 to bind to both the CE-LPH1 and SIF1 elements may be important in specifying intestinal expression of the corresponding genes. However, transcription factors in addition to Cdx-2 are probably required for the final developmental expression pattern of LPH and sucrase-isomaltase. Whereas LPH activity is high at birth and is down-regulated in most mammals after weaning, sucrase-isomaltase levels are low at birth but increase significantly after birth.

Previous EMSA analysis suggests that Cdx-2 is an abundant CE-LPH1-binding protein in differentiated Caco-2 cells. Formation of the main CE-LPH1-protein complex with Caco-2 nuclear extract was weakened by antibodies to Cdx-2, although no supershifted complex was generated (10, 11). However, this does not preclude that other regulatory factors bind to CE-LPH1 under alternative binding conditions, or if a fractionated nuclear extract were used. The CE-LPH1 element contains two subregions which are both protected in DNase I-footprinting.
(5). Cdx-2 appears to bind to only the upstream TTTAC sub-region. We have suggested that a repressor binds to the other half-site because a mutation here, which does not affect Cdx-2 binding, leads to higher transcriptional activity of a linked reporter gene (10).

A one-hybrid genetic screen in yeast (12) was carried out to identify cDNAs from Caco-2 cells, which encode factors able to bind specifically to the CE-LPH1 element. This approach allows for the detection of DNA binding activity in vivo under physiological conditions, in contrast to EMSA analysis. DNA-binding factors that have previously been isolated using a one-hybrid screen include Olf1, an olfactory-specific transcription factor (12), ORC6, a protein that recognizes the yeast origin of DNA replication (13), and a metal response element-binding protein (14). Using this screen, a 1.65-kb cDNA was isolated, which was shown to contain the homeobox of the human HOXC11 protein, previously called HOX3H (15). We describe here the isolation of the full-length human HOXC11 sequence from HeLa cDNA using RACE (rapid amplification of cDNA ends). The relative affinity of HOXC11 protein for various DNA elements, the expression pattern of HOXC11 mRNA, and the ability of HOXC11 to activate transcription were analyzed.

EXPERIMENTAL PROCEDURES

Plasmid Construction—Standard molecular biology techniques were used (16), and sequencing of double-stranded DNA was carried out with ThermoSequenase (Amersham Pharmacia Biotech). All inserts generated by PCR (Advantage KlenTaq polymerase, CLONTech) were sequenced completely to check polymerase fidelity. Oligonucleotides were obtained from Amersham Pharmacia Biotech.

In order to construct a reporter plasmid driven by four CE-LPH1 elements, two copies of the annealed oligonucleotides 5′-GATCTTTTA-CAGTGTCTGTATACACCTGTTG-3′ and 5′-GATCTACTCTGAGTGTGGATTTAAA-3′ were cloned head-to-tail into pRS315HIS (kindly provided by R. Reed; Ref. 12) to prepare the reporter plasmid pRS-LPH. To construct the other reporter plasmids, annealed oligonucleotides containing BamHI and BglII overhangs were multimerized and cloned into BamHI/BglII-digested pBluescript, in which the SpeI site had first been replaced by a BglII site. The appropriate DNA fragments cut out with BamHI and SpeI were cloned into pRS315HIS. The plasmids pRS-LPHmut1, pRS-LPHmut2, pRS-LPHmut3, and pRS-LPHmut4 contained four head-to-tail copies of the mut1, mut2, mut3, and mut4 LPH oligonucleotides, respectively (Fig. 1). pRS-SIF contained two copies of the annealed oligonucleotides containing two SIF1 binding sites from the sarcomeric isomaltase gene promoter (17).

The chDc2 insert of pBAT7Cdx-3 (kindly provided by M. German; Ref. 9) was subcloned into the SaciEcoRI site of pBluescript and used as template for PCR with the universal M13 primer and a primer annealing to the I fragment of pF4 was subcloned into pBluescript and used as template for PCR with the universal M13 primer and a primer annealing to the I fragment of pF4. The 5′-GAATTCGACATTTTACTTCTGAGCTAC-3′. The 1-kb PCR product obtained was cloned into the SalI/EcoRI site of pUC86 (12), resulting in p86Cdx. The SalI/ApaI insert of p86Cdx was subcloned into pBluescript and used as template for PCR with the universal M13 primer and a primer annealing to the first in-frame methionine codon 5′-TAGCGGCCGCCCACCATGAAAAAGC-3′. The 600-bp PCR product was cloned into the NotI/NotI site of pRS315HIS (kindly provided by R. Reed; Ref. 12) to prepare the reporter plasmid pRS-LPHmut1, pRS-LPHmut2, pRS-LPHmut3, and pRS-LPHmut4 containing four head-to-tail copies of the mut1, mut2, mut3, and mut4 LPH oligonucleotides, respectively (Fig. 1). pRS-SIF contained two copies of the annealed oligonucleotides containing two SIF1 binding sites from the sarcomeric isomaltase gene promoter (17).

The double-stranded oligonucleotides used in this study are shown. The TTTA(T/C) core sequence is underlined. The bases that differ from the CE-LPH1–17 element are written in lowercase. BSK-17 represents a sequence present in the pRS315HIS vector upstream of the polylinker site that has extended homology to CE-LPH1; ABD-B-17 contains the optimal binding site for Abdominal-B, which was obtained by binding site selection (34).

The pGL3-LPH227WT plasmid was constructed by cloning the SalI fragment of pGL2-LPH227, which contains the pig LPH promoter region from −227 to −17, upstream of the luciferase reporter gene in pGL3-basic (Promega). The mutations that were generated by PCR overlap extension changed the TTTA(T/C) core sequence of CE-LPH1 and -1 to TTTA(T/C) in the plasmid pGL3-LPH227mut1ac, and the additional mutation of CE-LPH2ac from GTTC to GCCG. The size-fractionated cDNA was ligated into pCP86 (12) and electroporated into the Escherichia coli strain DH10B (Life Technologies, Inc.). The bacterial transformants (∼5 × 10⁶ with an average insert size of 1 kb) were scraped from the plates into LB medium, and plasmid DNA was prepared (Qiagen).

Yeast Methods—The yeast strain used throughout was yWAM2 (MATa Δpol4 URA3:GAL1-lacZ tlys2801amber his3-Δ200 trp1-Δ63 leu2 ade2–101 ochre CYH2; kindly provided by R. Reed; Ref. 12). Yeast and synthetic minimal media (SD—His) containing adenine sulfate, uracil, and lysine were prepared as described (21). Where appropriate, the minimal SD medium was supplemented with histidine, leucine, and/or tryptophan (Sigma) (21). Yeast transformed with the reporter plasmids were stably maintained in SD medium supplemented with tryptophan and histidine. Yeast were transformed by the polyethyleneglycol/LiAc method (22), and plasmid DNA was recovered from Amersham Pharmacia Biotech. Screening of the cDNA Library—Yeast containing the pRS-LPH plasmid were transformed with 6 μg of library cDNA and plated onto 6 SD—His 14-cm plates at an average density of 1.7 × 10⁶ transformants/plate. After 4 days of incubation at 30°C, 24 colonies were able to grow. These were colony-purified, and plasmid DNA was recovered and individually re-transformed into reporter plasmid-containing

*S. Nopsdberg, J. T. Troelsen, P. Carlsson, S. Enerbäck, H. Sjöström, and O. Norén, submitted for publication.
yeast. Transformant colonies (at least three independent ones for each) were resuspended in TE buffer at approximately 1000 viable cells/μl. 1 μl of each was plated onto selective media plates (SD + His or SD − His) and grown 2–4 days at 30 °C.

**Cloning the 5' End of HOXC11**—In order to clone the 5' end of HOXC11 cDNA, a modified RACE protocol was carried out using Marathon-Ready HeLa cDNA (CLONTECH) according to the manufacturer’s protocol except that Advantage Tth DNA polymerase (CLONTECH) was used. The gene-specific RACE primer used was F4PR1L (sequence 5'-GAAGACACGGCTGAAAGCTG-3', see Fig. 3). The reaction mix was overlaid with oil and cycled in a Hybaid thermal cycler where the tube temperature was programmed for 1 cycle of 94 °C for 1 min, 5 cycles of 94 °C for 30 s and 72 °C for 4 min; 35 cycles of 94 °C for 30 s and 70 °C for 4 min; and 1 cycle of 70 °C for 5 min. A portion of the reaction was loaded onto an agarose gel and analyzed by Southern blotting with a 410-bp HincII-digested fragment of 92B1, which was radioactively labeled using the Rediprime kit (Amersham Pharmacia Biotech). The 550-bp RACE product thus identified was cut out and cloned (TA cloning kit, Invitrogen).

**Cell Culture, Transfections, Preparation of Whole Cell Extracts, and EMSAs**—Caco-2 and HeLa cells were grown in minimum essential medium (Life Technologies, Inc.) containing 10% calf serum. Caco-2 cells were seeded at 2 × 104 cells/cm² flask, and were harvested after 4 days (80% confluent) and 14 days (1 week post-confluence). HeLa cells were transfected by calcium phosphate precipitation with either pRC-HOXAN or water (untransfected control) as described previously (10). Whole cell extracts were prepared 48 h later using the protocol of Scholer et al. (24). Protein concentration was determined using Bradford reagents (Sigma Diagnostics). EMSAs were carried out as described previously (10).

Transient transfections of Caco-2 cells by the calcium phosphate method were carried out as described previously (10). All transfections contained 1 μg of a β-galactosidase expression plasmid as an internal standard for transfection efficiency (PCH110; Amersham Pharmacia Biotech) and 2.5 μg of pGL3-LPH227 wild-type or mutated constructs. In addition, 1.25 μg of pRC-CMV, pRC-HOXAN, pRC-HOX, and/or 0.63 μg of RSV-HNF, made up to 2.5 μg with pRC-CMV, were co-transfected.

**Analysis of RNA Expression**—Northern analysis was carried out according to standard techniques (25) using 2 μg each of poly(A)+ RNA from HeLa cells and small intestine (CLONTECH) and undifferentiated (80% confluent) and differentiated (1 week post-confluent) Caco-2 cells. Poly(A)+ RNA was prepared from Caco-2 cells using the Fast-Track kit (Invitrogen). A 400-bp Radioactive riboprobe (Ambion) was labeled to low specific activity (7 × 10⁴ cpm/ml Express Hyb solution (CLONTECH) was carried out according to the manufacturer’s protocol.

For ribonuclease protection assay, the Promega T3/T7 polymerase kit was used to synthesize antisense riboprobes from linearized p2B410 and pLPH380 plasmid templates to high specific activity (1 × 10⁶ cpm/ml). Both probes were gel-purified according to the protocol provided with the Ambion RAPIDII kit. The cyclin-dependent kinase control antisense riboprobe (Ambion) was labeled to low specific activity (7 × 10⁵ cpm/μg) and purified by RNAid (Bio101). 1 μg of yeast tRNA or polyclonal RNA, each supplemented with 10 μg of yeast tRNA, was hybridized overnight at 50 °C with 10⁶ cpm HOXC11 riboprobe and a mixture of 3 × 10⁶ cpm LPH riboprobe and 3000 cpm cyclin-dependent kinase control riboprobe and analyzed after RNA digestion on a non-denaturing polyacrylamide gel, according to the Ambion protocol, together with radioactively labeled double-stranded DNA markers.

PCR with double-stranded cDNA templates was carried out with 10 pmol each of the following primer pairs: glyceraldehyde-3-phosphate dehydrogenase (G3PDH) (5'-GACCAAGCTTCAAGAGCAG-3' and 5'-TCCAACACCTGTTGCATAG-3'), LPH (5'-GGAGGTTACAGTTTC-3' and 5'-GAGGATTGGCCAGGCTG-3'), or HOX11 (5'-AACTTTTCTCGCCATCTCGTT-3' and 5'-GCAGCAAGACATTGTCGCCGAGGTG-3'). The reaction mix was overlaid with oil and cycled in a Hybaid thermal cycler where the tube temperature was programmed for 1 cycle of 94 °C for 1 min, 4 cycles of 94 °C for 30 s and 72 °C for 4 min; 35 cycles of 94 °C for 30 s and 70 °C for 4 min; and 1 cycle of 70 °C for 5 min. A portion of the reaction was loaded onto an agarose gel and analyzed by Southern blotting with a 410-bp HincII-digested fragment of 92B1, which was radioactively labeled using the Rediprime kit (Amersham Pharmacia Biotech).

**RESULTS**

**Cloning of a CE-LPH1-binding Protein**—In order to clone a factor capable of binding to the CE-LPH1 element of the LPH gene promoter, the CE-LPH1 element was multimerized and placed in front of a HIS3 reporter gene in the construct pRS-LPH. Yeast containing the pRS-LPH reporter plasmid were transfected with a cDNA library prepared from differentiated Caco-2 cells, in which cDNA-encoded proteins are expressed as a fusion (hybrid) with the activation domain of the yeast transcriptional activator GAL4 (Fig. 2A). All yeast co-transformants grow on minimal medium supplemented with histidine (SD + His); the ability to grow on medium lacking histidine (SD - His) correlates with the extent of HIS3 reporter gene activation by the GAL4 fusion protein. The plasmid pPC86, which lacks a DNA insert, is unable to increase activation of the HIS3 gene, so there is no growth on SD - His (Fig. 2B).

Out of 10⁶ library transformants, 24 were able to grow in the absence of histidine. Two of these grew slowly on selective media, contained insert-less plasmid DNA, and were not studied further. An additional 20 clones encoded proteins that activated the reporter gene equally well in the presence (pRS-LPH) or absence of CE-LPH1 sites (pRS315HIS and p59.7). The p59.7 plasmid contains three copies of an unrelated se-
quence element, which binds an olfactory transcription factor, Olf-1 (12). The significant number of false positives are presumed to encode proteins that bind DNA nonspecifically. The remaining two yeast positives, pF4 and pF6, did not activate from a reporter gene driven by heterologous Olf-1 binding sites (p59.7). In addition, only the two true positives (pF4 and pF6) were able to grow under increased stringency conditions, i.e. in the presence of 5 mM 3-aminotriazole (Sigma; data not shown). Both plasmids contained identical 1.65-kb cDNA inserts, as judged by restriction digest. The insert of pF4 was sequenced completely on both strands, and shown to contain an open reading frame of 194 amino acids including the homeodomain of human HOXC11, in-frame with the vector-encoded GAL4 activation domain (Fig. 3). The isolated pF4 plasmid DNA was re-transformed into yeast containing either the pRS-LPH reporter plasmid used in the library screen or other constructs designed to test the specificity of binding (Fig. 2B). For comparison, yeast were also transformed with the plasmid p86Cdx, which encodes a fusion protein of the activation domain of GAL4 with full-length shCdx-2 (9). Cdx-2 has previously been shown to bind to the CE-LPH1 element (10). As shown in Fig. 2B, both the p86Cdx and pF4-encoded fusion proteins activate HIS3 reporter gene expression when driven by SIF1, CE-LPH1, or mut2 LPH1 elements, resulting in growth on SD - His plates. These elements also bind NF-LPH1 from Caco-2 cells (6). On the other hand, there is no activation, or reduced activation, from mut1 LPH1, mut3 LPH1, and p59.7-driven reporter genes. The mut1 and mut4 LPH1 oligonucleotides do not bind NF-LPH1, whereas mut3 LPH1 binds weakly, as shown by competition EMSA analysis (6). Although the parental pRS315HIS reporter plasmid lacks CE-LPH1 elements, there

![Image](https://via.placeholder.com/150)
is some activation of the HIS3 gene by p86Cdx and pF4, resulting in weak growth on minimal medium lacking histidine (Fig. 2B). This is probably due to the presence of a CE-LPH1–like sequence just upstream of the polylinker. A double-stranded oligonucleotide covering this vector sequence, referred to as BSK-17 in Fig. 1, was able to compete for HOXC11 binding to the CE-LPH1–24 probe in EMSA analysis (Fig. 5, lane 9).

Isolation of Full-length HOXC11 DNA—Hox genes in man and mouse are organized in four chromosomal clusters, which have probably arisen during evolution by gene duplications (reviewed in Ref. 26). It was expected that HOXC11 would show greatest sequence homology with the human HOXA11 and HOXD11 paralogues, which are located at a similar position on the A and D chromosomal clusters; there is no corresponding transcript on the B cluster (27). A putative sequence entry for the N terminus of human HOXC11 was identified by a data base search with the N-terminal sequence of mouse Hoxa11. The corresponding plasmid 92B1 (accession number Z63886) was acquired from the I.M.A.G.E. Consortium (18) and sequenced. The 92B1 genomic insert contains the first 351 nucleotides of the pF4 cDNA insert flanked on the 5′ side by 325 additional nucleotides and on the 3′ side by divergent, presumably intronic, sequence (Fig. 3). The presumed intron precedes the homeobox, and corresponds in its location to the intron present in the mouse Hoxa11 and Hoxd11 genes (Fig. 4).

In order to obtain the complete 5′ nucleotide sequence of HOXC11, RACE was carried out using HeLa cDNA. HeLa cDNA was chosen because it gave a strong PCR product in PCR analysis using HOXC11-specific primers (Fig. 9). The RACE fragment was cloned, and five independent clones were sequenced. The sequence of the longest RACE product is shown in Fig. 3 and is identical to the genomic sequence obtained with the Genome Walker kit from CLONTECH (data not shown). The first methionine of the longest cDNA is in a good translational context (28) and is preceded in the genomic sequence by an in-frame stop codon (data not shown). The deduced full-length cDNA is 2.0 kb and includes an open reading frame, which encodes a protein with 304 amino acids.

As shown in Fig. 4A, there are scattered regions of homology between human HOXC11 and mouse Hoxa11 and Hoxd11 outside the homeodomain. A glycine/alanine-rich sequence (38–45%) is also present in the three proteins just before the homeodomain. Similar regions in transcription factors have been implicated in both transcriptional repression (29) and activation (30). The homeodomain and C-terminal extension of human HOXC11 are identical to that of mouse Hoxc11 (31) and are also highly conserved with the mouse paralogous group.
proteins, Hoxa11 and Hoxd11 (32, 33) (Fig. 4B). Conservation between the homeodomains of *Drosophila* Abdominal-B and mammalian HOXC11, Hoxa11, and Hoxd11 is especially evident in the recognition helix 3 and N-terminal arm, both of which contact DNA.

**Characterization of HOXC11 DNA Binding Activity**—In order to analyze HOXC11 DNA binding activity, expression of the pF4 cDNA insert was put under the control of a cytomegalovirus promoter. The resulting plasmid pRC-HOXCAN, encoding the C-terminal 190 amino acids of HOXC11, was transfected into HeLa cells, which do not contain an endogenous CE-LPH1 binding activity (10). No complex is formed using whole-cell extract from untransfected HeLa cells in EMSA analysis (Fig. 5, lane 1), whereas a specific complex is formed with extract from transfected HeLa cells (Fig. 5, lane 2). HOXC11 is expected to bind to similar DNA sequences as the related *Drosophila* Abdominal-B protein. Therefore, the binding affinity of HOXC11 to CE-LPH1 and to an optimized Abdominal-B binding site was compared. The CE-LPH1-protein complex formed with extract from pRC-HOXCAN transfected HeLa cells is competed to a similar extent by cold competitor CE-LPH1–17 (Fig. 5, lanes 6–8), a related probe which has three base changes to include the optimal binding site for Abdominal-B (34). An unspecific oligonucleotide did not compete (Fig. 5, lane 10).

The ability of elements present in the yeast reporter plasmids shown in Fig. 2B to bind HOXC11 was analyzed by competition using either CE-LPH1–24 (Fig. 6A) or SIF1–24 (Fig. 6B) as probe in EMSA analysis. The binding affinity of HOXC11 appears to be highest for SIF1–24 and CE-LPH1–24 elements, intermediate for CE-LPH1–17 and mut2 LPH elements, and low for mut1 LPH, mut3 LPH, mut4 LPH, and unspecific elements. This result is consistent with EMSAs carried out with the endogenous NF-LPH1 present in Caco-2 nuclear extracts (6).

**Pattern of HOXC11 Expression**—We used Northern blot analysis to examine the expression and size of the HOXC11 mRNA transcript (Fig. 7). A probe from the 3'-untranslated region of HOXC11 was used to compare levels of HOXC11 expression in HeLa cells, Caco-2 cells in their undifferentiated (80% confluent) and differentiated (1 week post-confluent) state, and small intestine. One-week post-confluent Caco-2 cells have started to express markers of intestinal differentiation, including LPH (7) (Figs. 8 and 9).

HeLa cells contain a substantial amount of a 2.1-kb transcript that hybridizes to the HOXC11 probe. This correlates well in size with the full-length HOXC11 cDNA determined by RACE on HeLa cDNA (2.0 kb, Fig. 3). An additional transcript of 1.7 kb, corresponding to that reported for a hematopoietic cell line (35), is found in HeLa and also in Caco-2 cells but not in the small intestine from a 15-year old (Fig. 7, top panel). The amount of this transcript increases upon differentiation of Caco-2 cells, as judged by hybridization of the same filter with an actin probe (Fig. 7, lower panel).

The same mRNAs used in Northern blot analysis were analyzed by ribonuclease protection assay to determine the 5’ end of HOXC11 mRNA (Fig. 8). An antisense riboprobe correspond-
Kodak MS film at 80% confluent Caco-2 cells. In the lower panel, the membrane was hybridized with a DNA fragment from the 3’ untranslated region of HOXC11; the positions of the strongest hybridizing bands are indicated. Caco-2 UD refers to undifferentiated (80% confluent) Caco-2 cells, Caco-2 Diff. refers to differentiated (1 week post-confluent) Caco-2 cells. In the lower panel, the membrane was hybridized with an actin probe. The membrane was exposed to Biomax Kodak MS film at –70 °C for 3 days (top panel) or 2 h (lower panel). The gel was blotted onto a Hybond N+ membrane. In the top panel, the membrane was hybridized with a DNA fragment from the 3’-untranslated region of HOXC11; the positions of the strongest hybridizing bands are indicated. Caco-2 UD refers to undifferentiated (80% confluent) Caco-2 cells, Caco-2 Diff. refers to differentiated (1 week post-confluent) Caco-2 cells. In the lower panel, the membrane was hybridized with an actin probe. The membrane was exposed to Biomax Kodak MS film at –70 °C for 3 days (top panel) or 2 h (lower panel).

**FIG. 7.** Northern blot analysis of HOXC11 expression. The indicated poly(A)+ RNA (2 µg each) were separated on an agarose/formaldehyde gel. The positions of the 18 S and 28 S ribosomal particles were determined by loading 2 µg of total RNA in separate lanes on the same gel. The gel was blotted onto a Hybond N+ membrane. In the top panel, the membrane was hybridized with a DNA fragment from the 3’-untranslated region of HOXC11; the positions of the strongest hybridizing bands are indicated. Caco-2 UD refers to undifferentiated (80% confluent) Caco-2 cells, Caco-2 Diff. refers to differentiated (1 week post-confluent) Caco-2 cells. In the lower panel, the membrane was hybridized with an actin probe. The membrane was exposed to Biomax Kodak MS film at –70 °C for 3 days (top panel) or 2 h (lower panel).

**FIG. 8.** Analysis of HOXC11 and LPH expression by ribonuclease protection. In lanes 1–6, 1 µg of yeast transfer RNA (tRNA) or the indicated poly(A)+ RNA was hybridized overnight with a riboprobe corresponding to the 5’ region of HOXC11, subjected to RNase digestion, and loaded onto a non-denaturing acrylamide gel. In lanes 7–12, 1 µg each of the indicated RNA was hybridized overnight with a mixture of two riboprobes corresponding to the 280-bp BamHI/Pst1 fragment from the LPH gene and the 103-bp fragment of the cyclophilin gene. The positions of the main protected double-stranded RNAs are indicated by arrows. The bands marked with asterisks (*) are present in the tRNA control lanes and probably correspond to undigested probe. The position of dsDNA marker bands are indicated. The dried gel was exposed to Biomax Kodak MS film at –70 °C for 2 days.

**FIG. 8.** Analysis of HOXC11 and LPH expression by ribonuclease protection. In lanes 1–6, 1 µg of yeast transfer RNA (tRNA) or the indicated poly(A)+ RNA was hybridized overnight with a riboprobe corresponding to the 5’ region of HOXC11, subjected to RNase digestion, and loaded onto a non-denaturing acrylamide gel. In lanes 7–12, 1 µg each of the indicated RNA was hybridized overnight with a mixture of two riboprobes corresponding to the 280-bp BamHI/Pst1 fragment from the LPH gene and the 103-bp fragment of the cyclophilin gene. The positions of the main protected double-stranded RNAs are indicated by arrows. The bands marked with asterisks (*) are present in the tRNA control lanes and probably correspond to undigested probe. The position of dsDNA marker bands are indicated. The dried gel was exposed to Biomax Kodak MS film at –70 °C for 2 days.

**FIG. 9.** A. PCR analysis of HOXC11 and LPH expression. Each of the indicated cDNAs was subjected to PCR using the indicated primer set. The PCR was carried out in a Robocycler (Stratagene) for 1 min at 94 °C, followed by 5 cycles of 30 s at 94 °C and 3 min at 72 °C, followed by the indicated number of cycles of 30 s at 94 °C and 3 min at 70 °C. The G3FDH primers amplify a 450-bp fragment of the human G3FDH housekeeping gene. The LPH and HOXC11 primers amplify a 400-bp and a 690-bp fragment, respectively, from regions of the LPH and HOXC11 cDNAs which contain an intron in the genomic sequence. The PCR primers were designed to bind to regions of HOXC11, which are poorly conserved in other HOX gene products. M, 100-bp marker. B, PCR analysis of HOXC11 in human fetal tissues. Each of the indicated cDNAs was subjected to PCR as above, using G3FDH or HOXC11 primer sets. In the bottom panel, a Southern blot of the PCR reactions generated with the HOXC11 primer set was probed with a radioactively labeled HOXC11-specific oligonucleotide. F, fetal.

**FIG. 9.** A. PCR analysis of HOXC11 and LPH expression. Each of the indicated cDNAs was subjected to PCR using the indicated primer set. The PCR was carried out in a Robocycler (Stratagene) for 1 min at 94 °C, followed by 5 cycles of 30 s at 94 °C and 3 min at 72 °C, followed by the indicated number of cycles of 30 s at 94 °C and 3 min at 70 °C. The G3FDH primers amplify a 450-bp fragment of the human G3FDH housekeeping gene. The LPH and HOXC11 primers amplify a 400-bp and a 690-bp fragment, respectively, from regions of the LPH and HOXC11 cDNAs which contain an intron in the genomic sequence. The PCR primers were designed to bind to regions of HOXC11, which are poorly conserved in other HOX gene products. M, 100-bp marker. B, PCR analysis of HOXC11 in human fetal tissues. Each of the indicated cDNAs was subjected to PCR as above, using G3FDH or HOXC11 primer sets. In the bottom panel, a Southern blot of the PCR reactions generated with the HOXC11 primer set was probed with a radioactively labeled HOXC11-specific oligonucleotide. F, fetal.

**HOXC11 Stimulates HNF1α-dependent Transcription from the LPH Promoter**—We have previously demonstrated that the transcription factor HNF1α binds to three sites in the 894-bp upstream flanking region of the LPH gene. The proximal 227 bp of the LPH gene promoter is sufficient to drive differentiation-
stimulation of transcription, respectively. In the case of pGL3-N or pRC-HOX produces a further 7-fold or 19-fold pRC-HOX D activates the wild-type LPH promoter 2-fold, co-transfection of in LPH and Sucrase-Isomaltase (SI)—We have previously dem-

dependent transcription of a reporter gene in differentiated Caco-2 cells, to a level 80% of that seen with the 894-bp LPH dependent transcription by HOXC11 is much reduced for pGL3-LPH227mut1ac,2c in which the potential binding sites for

FIG. 10. HOXC11 stimulates HNF1α-dependent activation of the LPH promoter. Subconfluent Caco-2 cells were transfected with pGL3-LPH227WT, which contains the wild-type LPH promoter from position -227 to -17 in front of the luciferase gene, with or without the HNF1α expression plasmid RSV-HNF. The LPH227mut1ac construct is mutated in CE-LPH1a and -1c, and the LPH227mut1ac,2c construct has an additional mutation in the HNF1α-binding site, CE-LPH2c. The effect of co-transfecting expression plasmids for the short or long form of HOXC11 was tested. Luciferase activity was corrected for variation in transfection efficiency and is presented as fold-activation relative to pGL3-LPH227WT alone. The means ± S.D. were calculated from three experiments.

HOXC11 and Cdx-2 are both homeodomain proteins. The homeodomain is a 60-amino acid stretch, which contains a helix-turn-helix motif that is involved in DNA recognition. Many homeodomain proteins bind preferentially to a TAAT core sequence. HOXC11 is one of 15 mammalian Hox genes, which are related to the Drosophila homeotic gene for Abdominal-B. The Abd-B-like Hox proteins share several distinguishing features; they are expressed most posteriorly, their homeodomains recognize a divergent TTAT sequence (34), they do not contain a pentapeptide motif conserved in other Hox genes, and they differ in their ability to interact with Pbx proteins (36, 37). We have demonstrated similar affinity of HOXC11 for an optimal Abd-B sequence as for the CE-LPH1 element, which contains a related core sequence TTAC. Two residues in the N-terminal arm of the Abd-B homeodomain, K6 and P7, appear to play an essential role in the divergent specificity of Abd-B compared with other Drosophila homeoproteins (34). The N-terminal arm of HOXC11 conserves one of these critical residues, the proline at position 7.

It has recently been demonstrated by binding site selection that the Abd-B-like Hox proteins Hoxb-9, Hoxa-11, Hoxd-12, and Hox-13 bind preferentially to a TTTAC-containing sequence (37). However, not much is presently known about the natural binding sites for mammalian Abd-B-type Hox proteins. The few cases where target genes have been identified involve auto- and cross-regulatory interactions between different Hox genes and their gene products. In the case of the Abd-B-type proteins, HOXD9 and HOXD10 (previously called HOX4C and HOX4D, respectively) have been shown to trans-activate the gene for HOXC5 in cotransfection studies, mediated by a 5′-AATTTATGA-3′ sequence in the HOXC5 promoter (38). In the paralogous mouse Hoxa5 and Hoxb5 gene promoters, the corresponding element resembles the CE-LPH1 element: 5′-AATTTACGAC-3′. Thus, Abd-B-like Hox proteins recognize a

DISCUSSION

HOXC11 and Cdx-2 Bind to Proximal Regulatory Elements in LPH and Sucrase-Isomaltase (SI)—We have previously dem-

onstrated that 1 kb of the lactase gene promoter is sufficient to direct small intestine-specific expression and post-weaning decline in transgenic mice (3). We have focused our attention on a proximal promoter element, CE-LPH1, that is bound by an intestine-specific factor which co-varies with lactase expression (5). Deletion of the CE-LPH1 element in stably transfected Caco-2 cells leads to a 64% reduction in LPH promoter activity in differentiated cells. Using a yeast genetic screen to identify factors capable of binding to CE-LPH1, the HOXC11 cDNA was isolated twice in a screen of 10^6 clones. This suggests that HOXC11 is more abundantly expressed in differentiated Caco-2 cells and/or binds the CE-LPH1 element with higher affinity, compared with other DNA-binding proteins encoded in the cDNA library. However, it should be noted that the cDNA used in the library construction was subjected to a complete digestion with NotI, resulting in the probable exclusion of some cDNAs.

Previously, a factor that binds to the related SIF1 element in the gene for intestinal SI was isolated and identified as the homeodomain protein Cdx-2 (8). In our present work, we have shown that both HOXC11 and Cdx-2, expressed from a single-copy plasmid in yeast cells, are able to bind to the CE-LPH1 and SIF1 elements under in vivo conditions, in a manner dependent on an intact TTTA(T/C) sequence. Mutation of bases within this 5-nucleotide site, as in mut1 LPH1 and mut4 LPH1, reduces binding of HOXC11 both in the yeast assay and in EMSA analysis. Although binding of HOXC11 to mut3 LPH1 is intermediate between that of mut1 LPH1 and mut2 LPH1, binding of the factors to mut3 LPH1 in yeast does not lead to growth on minimal media, due perhaps to the all-or-none nature of the yeast screen.

HOXC11 and Cdx-2 are both homeodomain proteins. The homeodomain is a 60-amino acid stretch, which contains a helix-turn-helix motif that is involved in DNA recognition. Many homeodomain proteins bind preferentially to a TAAT core sequence. HOXC11 is one of 15 mammalian Hox genes, which are related to the Drosophila homeotic gene for Abdomi-
TTAT(C)/C core sequence, which is likely to be physiologically relevant.

There have been previous reports of intestinal factors other than Cdx-2 which bind to CE-LPH1 or SIF1 elements (20, 29, 39–41). For example, Traber et al. (41) detected a SIF1-binding protein in a fractionated Colo-DM extract that had a lower molecular weight than Cdx-2. We suggest that HOXC11 may interact directly with other and whether they co-operate in DNA binding.

It is possible that the intestinal specific expression of LPH is dependent, at least in part, on the overlap in tissue distribution of both HNF1α and HOXC11. The presence of HOXC11 in tissues outside the intestine, such as HeLa cells, fetal kidney, and fetal skeletal muscle, suggests that it has additional functions. Homeobox genes typically play an important role in tissue development; for example, pdx-1 (pancreatic and duodenum homeobox gene 1) is essential for pancreatic development in mice (47). Diverse mammalian Abd-B-type HOX genes have an important role in limb and urogenital development (32, 48–50).

The expression of HOXC11 in human fetal intestine suggests that HOXC11 plays an important role in intestinal development.

Acknowledgments—We thank Liselotte Laustsen and Anette Melchior Jensen for excellent technical assistance and Dr. Jurgen Olsen and Dr. Nikolaj Spodsberg for stimulating discussions. We are also grateful to Dr. R. Reed for the plasmids pPC86, p59.7, pRS531HHS, and y11 and yeast strain yWAM2; Dr. M. German for providing us with the plasmid pBAT7Cdxx-3; and Dr. F. Tronche for the plasmid RSV-HNF1α.

REFERENCES

1. Krasinski, S. D., Estrada, G., Yeh, K. Y., Yeh, M., Traber, P. G., Rings, E. H., Buller, H. A., Verhave, M., Montgomery, R. K., and Grand, R. J. (1994) Am. J. Pathol. 144, 455–469

2. Lloyd, M., Merissen, G., Fischer, M., Olsen, W., Goosspeed, D., Gemini, M., Boll, W., Semenza, G., and Mantei, N. (1992) J. Clin. Invest. 89, 524–529

3. Troelsen, J. T., Mehlem, A., Olsen, J., Spodsberg, N., Hansen, G. H., Prydz, H., Noreen, O., and Sjöstrom, H. (1984) FASEB J. 24, 2583–2585

4. Montgomery, R. K., Buller, H. A., Rings, E. H., and Grand, R. J. (1993) FASEB J. 7, 2824–2832

5. Troelsen, J. T., Olsen, J., Noreen, O., and Sjöstrom, H. (1992) J. Biol. Chem. 267, 9207–9210

6. Troelsen, J. T., Olsen, J., Mitchelmore, C., Hansen, G. H., Sjöstrom, H., and Noreen, O. (1994) FASEB J. 8, 297–301

7. Chantret, I., Barbat, A., Dausseuf, A., Brattain, M. G., and Zweibaum, A. (1988) Cancer Res. 48, 1356–1362

8. Suh, E., and Traber, P. G. (1996) Mol. Cell. Biol. 16, 619–625

9. German, M. S., Wang, J., Chadwick, R. B., and Rutter, W. J. (1992) Genes Dev. 6, 2165–2176

10. Troelsen, J. T., Mitchelmore, C., Spodsburg, N., Jensen, A. M., Noreen, O., and Sjöstrom, H. (1997) Biochem. J. 322, 833–838

11. Olsen, J., Lloyd, M., Korsmo, H., and He, Y. Z. (1996) Am. J. Pathol. 147, 477–481

12. Wang, M. M., and Reed, B. R. (1993) Nature 364, 121–126

13. Li, J. J., and Herskowitz, I. (1995) Science 263, 1870–1874

14. Inouye, C., Remondelli, P., Karin, M., and Elledge, S. (1994) DNA Cell Biol. 13, 731–742

15. Stornaiuolo, A., Aicamora, D., Punanse, M., D’Esposito, M., Morelli, F., Migliacchi, E., Rambaldi, M., Paella, A., Nigro, V., and Simeone, A. (1996) Cell Differ. Cell. Dev. 31, 119–127

16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY

17. Traber, P. G., Wu, G. D., and Wang, W. (1992) Mol. Cell. Biol. 12, 3614–3627

18. Lennon, A., Auffray, C., Polymeropoulos, M., and Szafer, M. B. (1996) Genomics 33, 150–159

19. Mantei, N., Villal, M., Enderle, T., Wacker, H., Boll, W., Semenza, G., and Mantei, N. (1992) EMBO J. 11, 2705–2713

20. Chomczynski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156–159

21. Sherman, F. (1991) Methods Enzymol. 194, 3–21

22. Gietz, D., St. Jean, A., Woods, R. A., and Schiestl, R. H. (1992) Nucleic Acids Res. 20, 4125

23. Hoffman, C. S., and Winston, F. (1987) Genet. 57, 267–272

24. Scholer, H. R., Hatzopoulos, A. K., Balling, R., Suzuki, N., and Gruss, P. (1989) EMBO J. 8, 2543–2550

25. Ausioel, F. M., Brett, R., Kingstom, B. E., Moore, D. D., Seidman, J. G., Smith, J., and Struhl, K. (1997) Current Protocols in Molecular Biology, John Wiley & Sons, New York

26. Gehring, W. J., Afolter, M., and Burglin, T. (1994) Annu. Rev. Biochem. 63, 487–526

27. Scott, M. P. (1993) Nucleic Acids Res. 21, 1687–1688

28. Cavener, D. R., and Ray, S. C. (1991) Nucleic Acids Res. 19, 3185–3192

29. Catron, K. M., Zhang, H., Marshall, S. C., Inostroza, J. A., Wilson, J. M., and Abate, C. (1995) Mol. Cell. Biol. 15, 861–871

30. Meijer, D., Graus, N., and Grosveld, G. (1992) Nucleic Acids Res. 20, 2241–2247

31. Peterson, R. L., Papenbrock, T., Davda, M. M., and Awgulewitsch, A. (1994) Mech. Dev. 47, 253–260

32. Hsieh-Li, H. M., Witte, D. P., Weinstein, M., Branford, H. L., Small, K., and Potter, S. S. (1997) Development 123, 1373–1385

33. Izsipus-Belmonte, J. C., Falkenstein, H., Dolle, P., Renucci, A., and Duboule, D. (1991) EMBO J. 10, 2279–2289

34. Ekkert, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E., and
Beachy, P. A. (1994) *EMBO J.* **13**, 3551–3560
35. Magli, M. C., Barba, P., Celetti, A., De Vita, G., Cillo, C., and Boncinelli, E. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6348–6352
36. Phelan, M. L., Rambaldi, L., and Featherstone, M. S. (1995) *Mol. Cell. Biol.* **15**, 3899–3907
37. Shen, W.-F., Rosenfeld, S., Lawrence, H. J., and Largman, C. (1997) J. Biol. Chem. **272**, 8198–8206
38. Arcioni, L., Simeone, A., Guazzi, S., Zappavigna, V., Boncinelli, E., and Mavilio, F. (1992) *EMBO J.* **11**, 265–277
39. Hecht, A., Torbey, C. F., Kormo, H. A., and Olsen, W. A. (1997) Gastroenterology **112**, 803–812
40. Sibley, E., Santiago, N. A., and Gray, G. M. (1996) Gastroenterology **110**, A839
41. Suh, E., Chen, L., Taylor, J., and Traber, P. G. (1994) *Mol. Cell. Biol.* **14**, 7340–7351
42. Olsen, J., Laustsen, L., and Troelsen, J. T. (1994) FEBS Lett. **342**, 325–328
43. Wu, G. D., Chen, L., Forslund, K., and Traber, P. G. (1994) J. Biol. Chem. **269**, 17080–17085
44. Tronche, F., and Yaniv, M. (1992) BioEssays **14**, 579–587
45. Mendel, D. B., Khavari, P. A., Conley, P. B., Graves, M. K., Hansen, L. P., Admon, A., and Crabtree, G. R. (1991) Science **254**, 1762–1767
46. Rhee, K.-H., Stier, G., Becker, P. B., Suck, D., and Sandalzopoulois, R. (1997) J. Mol. Biol. **265**, 20–29
47. Jonsson, J., Carleson, L., Edlund, T., and Edlund, H. (1994) Nature **371**, 606–609
48. Mortlock, D. P., and Innis, J. W. (1997) Nat. Genet. **15**, 179–180
49. Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B. R. (1996) Science **272**, 548–551
50. Davis, A. P., Witte, D. P., Haib-Li, H. M., Potter, S. S., and Capecchi, M. R. (1995) Nature **375**, 791–795