Transcriptional regulation in endoderm development: characterization of an enhancer controlling Hnf3g expression by transgenosis and targeted mutagenesis

Holger Hiemisch, Günther Schütz and Klaus H. Kaestner

Molecular Biology of the Cell I Division, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany
and
Department of Genetics, University of Pennsylvania Medical School, 415 Curie Boulevard, Philadelphia, PA 19104-6145, USA

© Oxford University Press 1997

Introduction

Analysis of the regulation of liver-specific gene expression has led to the purification and cloning of several liver-enriched (but not liver-restricted) transcription factors. These factors include the following gene families (recently reviewed in Cereghini, 1996): the homeodomain-containing HNF-1 (hepatocyte nuclear factor 1) family, the winged helix domain-containing HNF-3 proteins, the orphan nuclear receptor family HNF-4, the C/EBP (CCAAT/enhancer binding proteins) basic leucine zipper proteins and the DBP family. Since all of these genes show a restricted expression pattern, the question arose as to how these transcription factors themselves are regulated. The first evidence for the existence of transcriptional cascades came from the study of differentiated and dedifferentiated hepatoma cell lines. A binding site for HNF-4 in the Hnf1a promoter has been shown to have a decisive role in the activation of Hnf1a, pointing to HNF-4 as a major activator of Hnf1a (Tian and Schibler, 1991; Kuo et al., 1992). Since the dedifferentiated cells used still contained HNF-1β and the HNF-3 proteins, it was suggested that they might be predetermined to the hepatic differentiation programme by these factors (see Cereghini, 1996). In fact, the Hnf3 genes are the transcription factors expressed earliest during definitive endoderm development (Ang et al., 1993; Monaghan et al., 1993; Sasaki and Hogan, 1993) and therefore might be involved in subsequent activation of other hepatic transcription factors. It was proposed that the HNF-3 proteins exert their regulatory role by remodelling the chromatin structure of the target genes in a way that makes it competent for subsequent activation events (McPherson et al., 1993; Gualdi et al., 1996; reviewed in Zaret, 1996). Gene targeting experiments aimed at elucidating putative roles of liver-enriched transcription factors (HNF-3β and HNF-4) in definitive endoderm formation failed to give conclusive answers because of early embryonic lethality (Ang and Rossant, 1994; Weinstein et al., 1994; Chen et al., 1994). In addition, the transcription factor HNF-1α was shown to be dispensible for early liver development through analysis of the mutant phenotype in Hnf1a−/− mice (Pontoglio et al., 1996). Thus, although expression of transcription factors at early stages of endoderm-derived organ formation is well documented, little is known about how these factors are activated and how they contribute to endoderm differentiation.

In order to better understand endoderm development we decided to study Hnf3g gene regulation in transgenic mice. We chose the Hnf3g gene for the following reasons. HNF-3γ is one of the transcription factors known to be expressed early during gut and liver formation and was suggested to play a role in regionalization of the gut (Monaghan et al., 1993). In contrast to the two other Hnf3 genes, Hnf3g is not expressed in the notochord and floor plate, therefore, we reasoned that the cis-regulatory elements governing its expression in the endoderm might be more accessible to analysis.

In the present study we show that an Hnf3g yeast artificial chromosome (YAC) targeted with a β-galactosidase reporter can faithfully recapitulate the endogenous expression pattern of Hnf3g. Guided by a DNase I hypersensitivity analysis of the gene locus, we were able to identify enhancers responsible for the expression pattern of the Hnf3g YAC. An enhancer of the Hnf3g gene driving expression in the posterior foregut and midgut endoderm...
Fig. 1. Targeting of a 170 kb YAC carrying the Hnf3g gene by insertion of a β-galactosidase reporter gene. (A) Structure of the unmodified Yγ5 YAC. Restriction sites for NotI are shown. The Hnf3g exons are represented by filled boxes and numbered. YAC vector arms are schematized with arrows (telomeres), open rectangles (yeast selective markers) and a crossed circle (autonomously replicating sequence and centromere). (B) Targeting scheme for replacement of exon 2 sequences with lacZ using a two-step procedure in yeast. The lacZ targeting construct is shown below the Hnf3g locus of the YAC harbouring the desired exon 2–lacZ fusion (2-lacZ) and a URA3 yeast selectable marker. The Bluescript vector backbone is indicated by a dashed line. A crossed line marks the recombination point of the pop in step between the YAC and the linearized targeting construct. Counter-selection against the presence of URA3 in the pop out step gives rise to the desired alteration of the gene locus. (C) PFGE with subsequent Southern blot hybridizations demonstrating the presence of the β-galactosidase gene in the Yγ5Z YAC. A control hybridization with an intronic probe for Hnf3g marks the positions of the parental (Yγ5) and the targeted YAC (Yγ5Z). The increase in size of Yγ5Z is due to the size difference between the β-galactosidase gene and the deleted portion of exon 2. λ, concatemeres of λ DNA (48.5 kb).

was shown to contain an HNF-1 binding site, which is essential for enhancer function in hepatoma cells, thus defining for the first time an HNF-1–HNF-3γ transcriptional cascade.

Results

Cloning of an Hnf3g–lacZ yeast artificial chromosome

Two Hnf3g YACs were isolated from a mouse YAC library and characterized (see Materials and methods). Briefly, hybridization patterns obtained with probes spanning the entire previously cloned mouse Hnf3g locus (Kaestner et al., 1994) revealed that the YACs contained un-rearranged inserts (data not shown). Restriction site mapping was used to locate the position of both exons of the Hnf3g gene within the YACs. A genomic map of the Yγ5 YAC, which was used for further manipulation, is shown in Figure 1A. In this YAC the Hnf3g gene is flanked by ~100 kb upstream and 60 kb downstream sequences. The YAC does not seem to have chimeric portions, since both insert ends were shown to stem from the same chromosomal origin (see Materials and methods). As we wanted to use the Hnf3g YAC for the analysis of cis-regulatory elements in transgenic mice, we needed to tag the YAC in order to differentiate its expression from that of the endogenous Hnf3g locus. We decided to tag the YAC with the bacterial gene encoding β-galactosidase using a pop in/pop out strategy (Figure 1B; for details see Materials and methods). The correct structure of the obtained YAC (Yγ5Z) was verified by several methods. Figure 1C shows pulsed field gel electrophoresis (PFGE) of Yγ5Z along with the parental Yγ5 YAC. Hybridizations to various probes demonstrate the correct integration of lacZ into the YAC.

Expression analysis of the Hnf3g–lacZ YAC in transgenic mice

DNA from YAC Yγ5Z was purified by PFGE as described previously (Schedl et al., 1993) and microinjected into FVB/N oocytes. Seven independent transgenic founders were obtained, five of which transmitted the transgene to their offspring. Three lines carried fragmented YACs judged by the absence of hybridization signals with probes for one or both vector arms (data not shown). The remaining lines carried two (line 5489) and four (line 5520) intact copies respectively and were analysed for Hnf3g/lacZ expression of the YAC transgene. Results of whole mount β-galactosidase staining of line 5520 are shown in Figure 2A and B, demonstrating transgene expression in the embryonic liver.

For a detailed analysis, embryos (E14.5) were sectioned and stained. As shown in Figure 2, lacZ expression exactly reproduced the endogenous expression pattern of Hnf3g.
Fig. 2. Embryonic β-galactosidase expression of the Hnf3g-lacZ YAC. Whole mount staining of E10.5 (A) and E12.5 (B) embryos (line 5520) showing liver expression. (C–H) Cryosections of E14.5 embryos (line 5520) revealing expression in (C) liver, (D) pancreas, (E) small intestine, (F) colon (Co), (G) ribs (Ri) and (H) epiphyseal cartilage (Ep) and actively proliferating cartilage of long bones (arrowhead). Bl, bladder. Bar corresponds to 0.5 mm in (A), 0.8 mm in (B), 25 μm in (C) and (D) and 100 μm in (E)–(H).

(Monaghan et al., 1993). Figure 2C and D shows transgene expression in liver and pancreas respectively. The patchy staining pattern in liver is very reminiscent of the results obtained by *in situ* hybridization (Monaghan et al., 1993) and probably reflects absence of HNF-3γ in blood islands. In the pancreas, staining is detected in cells that will form
the acini (arrowhead in Figure 2D). The YAC transgene shows a distinct pattern of expression along the anteroposterior axis of the developing gastrointestinal tract. Highest expression is seen in the mucosa of the stomach (data not shown) and developing colon (Co in Figure 2F). Weaker expression can be detected in the small intestine (Figure 2E), while no expression is found anterior to the stomach. In all parts of the digestive tract transgene expression is strictly restricted to the epithelial layer. Non-transgenic embryos did not show any background signals, thus verifying the specificity of the staining procedure (data not shown). Analysis of the YAC expression pattern allowed a more precise definition of Hnf3g expression in the developing bones than previously described. The transgene was expressed in ribs (Ri in Figure 2G), vertebrae (data not shown) and long bones. Figure 2H shows a parasagittal section through the humerus at the beginning of enchondral ossification. Strongest transgene expression was found in layers of actively proliferating cartilage (arrowhead), whereas no expression could be found in hypertrophic cartilage around the ossification centre. Staining in the epiphyseal cartilage is clearly visible (Ep in Figure 2H). There was no ectopic expression of the YAC transgene found in any of the embryos analysed. The staining pattern of line 5489 was qualitatively indistinguishable from that of the above described line 5520 (data not shown).

In order to obtain a quantitative measure of the expression levels of the two YAC transgenic lines, we isolated total RNA from a variety of adult organs. RNase protection assays were performed with probes that enabled us to directly compare transgene expression with the endogenous Hnf3g mRNA level (Figure 3A). Strongest YAC expression was found in stomach, colon and liver, whereas weaker signals were obtained from pancreas, jejunum and testis, very similar to the embryonic pattern. Signal intensities were quantified using a phosphorimager and calculated as YAC expression relative to one endogenous Hnf3g allele (Figure 3B). In all organs expression from the modified Hnf3g locus of the YAC was comparable with wild-type levels ranging from about one-quarter context (Kaestner et al., 1994). Interestingly, the four copy line 5489 produced roughly twice as much Hnf3g-lacZ mRNA as the two copy line 5489.

DNase I hypersensitive site analysis of the Hnf3g locus

To identify important cis-regulatory elements, we characterized the Hnf3g chromatin structure by comparing DNase I hypersensitivity in nuclei prepared from liver, where Hnf3g is expressed, and kidney, a non-expressing control tissue. Using a probe from the far 3′-region of Hnf3g, a set of DNase I hypersensitive sites (HS) were found to be specific for liver (+12 kb HS) or detected in both liver and kidney (+16 kb HS; Figure 4A). At the position of the promoter (P), a liver-specific HS was detected flanked by a non-specific HS close by. Another non-specific HS was found further upstream at −0.8 kb (data not shown). An additional liver-specific HS was identified in the intron (at +7.5 kb; not shown). The outermost HS found in both upstream and downstream regions of Hnf3g are organ-specific, flanking the inner liver-specific sites. A summary of all HS found in the Hnf3g locus is given in Figure 4B according to their position in the genomic context (Kaestner et al., 1994).

3′-Flanking sequences control expression of Hnf3g

Having established that the 170 kb around Hnf3g contained in YγZ is sufficient to reproduce expression of the endogenous gene in both embryos and adult animals, we were interested to dissect the relevant cis-elements. Guided by the DNase I HS identified above, we constructed a series of plasmid-derived transgenes comprising the promoter and 8 kb of 5′-flanking sequences together with either intronic or 3′-flanking sequences (Figure 5A). A summary of the embryonic expression pattern of the transgenic lines obtained by β-galactosidase staining of both whole mount embryos and cryosections is given in Figure 5B. Transgenics harbouring only the upstream region (pγZ) or these sequences together with the intron (pβ8Z) did not show Hnf3g-related expression. In contrast, transgenics derived from plasmid pγZ3.6 (which additionally contains sequences spanning the +12 kb HS), while showing extensive ectopic expression in all lines, exhibit the expected staining in developing bone in three out of five lines. When adding the sequence corresponding to the +16 kb HS (thus generating pβZ4.9) we were able...
Gene regulation of Hnf3g

Construct Transgenic line Transgene expression

<table>
<thead>
<tr>
<th>Construct</th>
<th>Transgenic line</th>
<th>Liver</th>
<th>Gut</th>
<th>Bone</th>
<th>Ectopic</th>
</tr>
</thead>
<tbody>
<tr>
<td>pγ8Z</td>
<td>4</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
<td>1/4</td>
</tr>
<tr>
<td>pγ8IZ</td>
<td>7</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>1/7</td>
</tr>
<tr>
<td>pγ8Z3.6</td>
<td>5</td>
<td>0/5</td>
<td>1/5</td>
<td>3/5</td>
<td>5/5</td>
</tr>
<tr>
<td>pγ8Z4.9</td>
<td>5</td>
<td>4/5</td>
<td>4/5</td>
<td>2/2</td>
<td>1/5</td>
</tr>
</tbody>
</table>

Fig. 5. Hnf3g plasmids with a β-galactosidase reporter gene used for transgenesis. (A) The different fragments containing the reporter gene alongside their genomic context and the HS found in liver are shown. lacZ, β-galactosidase gene and SV40 intron and poly(A); B, BamHI; E, EcoRI; H, HindIII.

enhancer element for the expression of Hnf3g in endoderm-derived tissues.

Deletion of sequences including the +12 kb and +16 kb HS abolishes expression of Hnf3g in liver, pancreas, stomach and small intestine

In the targeted mutation of Hnf3g, which deletes most of exon 2 plus 5.5 kb of 3′-non-translated sequence (Kaestner et al., unpublished results), the coding region was fused in-frame to a lacZ reporter gene. Surprisingly, no β-galactosidase expression was found in liver, pancreas, stomach, small intestine and developing bone (data not shown). Figure 7A shows a RNase protection assay with a probe that spans the Hnf3g–lacZ fusion and can thus be used to detect transcripts arising from the wild-type and mutant allele simultaneously. Hnf3g is expressed at ~50% of the wild-type levels in the samples obtained from the mutant allele simultaneously. Hnf3g–lacZ expression has been detected in small intestine (SI in Figure 6C), with lower levels in the umbilical hernia, in the colon (not shown) and developing pancreatic acini (P in Figure 6B). Moreover, YAC and plasmid expression followed the endogenous Hnf3g pattern in the nasal epithelium (Kaestner et al., unpublished results). Figure 6D presents evidence for β-galactosidase expression in ribs (Ri); staining in vertebrae and long bones is not shown. Ectopic expression in one out of five lines is seen in the epithelium of the lung (Figure 6D). In conclusion, ~13 kb of Hnf3g comprising pγ8Z4.9 contains the cis-regulatory elements necessary for all Hnf3g expression domains (liver, gut, pancreas, bone and nasal epithelium). Moreover, lack of expression in pγ8Z3.6 transgenics points to the additional sequence in pγ8Z4.9 as an important
Fig. 6. Embryonic β-galactosidase pattern of the plasmid transgene pγ8Z4.9 mimics expression of the Yγ5Z YAC. (A) Strong liver expression at E12.5 as revealed by whole mount staining. (B–D) Cryosections of E14.5 embryos of (B) liver (Li), pancreas (P) and stomach (St), (C) liver and small intestine (SI) and (D) ribs (Ri). The arrowhead in (B) points to the junction of squamous and glandular epithelium in the stomach. (D) Ectopic expression in the lung (Lu). Bar corresponds to 0.6 mm in (A), 200 μm in (B) and 100 μm in (C)–(D).

Fig. 7. The Hnf3g–lacZ allele is transcriptionally silent in liver, pancreas and stomach. (A) RNase protection analysis of 20 μg total RNA isolated from the tissues indicated from adult wild-type (+/+), heterozygous (+/−) and homozygous mutant (−/−) mice was performed with a probe which allows simultaneous detection of Hnf3g and Hnf3g–lacZ (labelled LacZ) mRNAs. The weak band observed in the −/− lanes for Hnf3g is non-specific, as it appears in the tRNA lane as well. (B) The chromatin structure of the entire Hnf3g locus is altered in livers of Hnf3g−/− mice. DNase I hypersensitive site analysis was performed with liver nuclei isolated from wild-type (+/+) or homozygous mutant (−/−) mice as described in Materials and methods. The locations of the HS near the promoter and at −0.8 kb are indicated.
wild-type and mutant litter mates were obtained and analysed as described in Materials and methods. As shown in Figure 7B, the HS site near the promoter of the gene is no longer present in the mutant samples, consistent with the inactive state of the gene in liver. Likewise, the HS site within the intron (at +7.5 kb) is lost in the −/− mice (data not shown). In contrast, the HS at −0.8 kb, which is present both in kidney and liver, was still present in the mutant animals. In summary, the DNase I HS pattern of the Hnf3g→− livers resembles that of wild-type kidney, a tissue where HNF-3γ is not expressed (compare Figure 4B). As indicated by DNase I hypersensitivity, deletion of the 3′-flanking region of Hnf3g has affected the chromatin structure >10 kb away, explaining the lack of expression of the Hnf3g→lacZ allele in liver and, by extension, in stomach, pancreas and small intestine. This finding identifies the deleted sequences 3′ of the HNF-3γ coding region as harbouring a dominant enhancer(s) directing expression in the posterior foregut and midgut endoderm.

Identification of a strong enhancer activity in the 3′-flanking region of Hnf3g

In order to define this posterior foregut and midgut endoderm enhancer activity in more detail, we cloned the relevant region in seven partially overlapping fragments in front of a Tk promoter/Cat reporter (pBLCAT5; Boshart et al., 1992) and transfected the plasmids into FTO-2B hepatoma cells, a cell line that had been shown previously to express high levels of Hnf3γ protein (Nitsch et al., 1993). As shown in Figure 8A, a strong enhancer activity was detected in a 1.3 kb BamHI fragment that corresponds to the +16 kb HS. Control transfections into Ltk− fibroblasts, which do not express HNF-3γ, revealed the cell specificity of the detected enhancer. A second fragment with minor activating potential was found within a 1.0 kb EcoRI–XbaI fragment including the sequences indicated by the +12 kb HS. In order to define the minimal sequences of the +16 kb enhancer, we transfected deletion derivatives of the 1.3 kb BamHI fragment into FTO-2B cells. After narrowing down the enhancer activity to a 700 bp DpnI fragment (data not shown), another deletion series of this fragment defined a 397 bp AluI fragment as necessary and sufficient for full activity (Figure 8B). The sequence is given in Figure 8C, together with putative transcription factor binding sites as determined by a computer search using binding site matrices (kindly performed by F.Tronche; Tronche et al., 1997). Among ubiquitous factors, a binding site for another liver-enriched transcription factor, HNF-1, was found in the +16 kb enhancer of Hnf3g.

HNF-1 is a potent transactivator of the +16 kb enhancer

In order to test the functional significance of the HNF-1 site in the +16 kb enhancer, we mutated the motif in the context of a 425 bp enhancer fragment in pBLCAT5 (Figure 9A) and transiently transfected this construct into FTO-2B hepatoma cells. Figure 9B demonstrates that mutation of the HNF-1 binding site results in a dramatic reduction in the +16 kb enhancer activity. The 15 bp motif labelled in Figure 9 has independently been used by others in band shifts with liver extracts and was shown to possess high binding affinity for HNF-1 (Tronche et al., 1994). It remains to be tested whether HNF-1 can transactivate the +16 kb enhancer. The HNF-1 family of homeodomain proteins consists of HNF-1α and HNF-1β (for a review see Tronche et al., 1994). C33 human cervical carcinoma cells thus were transiently transfected with increasing amounts of HNF-1 expression plasmid (HNF-1α and HNF-1β, driven by a RSV promoter, kindly provided by F.Tronche; Chouard et al., 1990; Rey Campos et al., 1991), together with either the 425 bp wild-type enhancer or the enhancer with the mutated HNF-1 binding site (Figure 9C). The upper panel shows that HNF-1α is able to stimulate the wild-type enhancer but not the mutated form. HNF-1β displays a very similar transactivating potential, as demonstrated in the lower part of Figure 9C. C33 cells do not show a decreased basal level of reporter gene activity upon mutation of the HNF-1 binding site, as shown above for FTO-2B cells, which can be explained by the absence of HNF-1 proteins in the cervical carcinoma cell line.

Discussion

A 170 kb YAC with an Hnf3g→lacZ fusion gives correct cell-specific and developmental expression

Studies in cell cultures have provided the basis for the understanding of cell type-specific regulation of gene expression in mammals. In many instances, however, the relevance of the postulated mechanisms for gene activity in the intact organism has not yet been demonstrated. In order to define the cis-elements required for Hnf3g gene expression in vivo we decided to start with YACs as a vector system that ensures accurate gene expression in transgenic mice (for a review see Lamb and Gearhart, 1995). In addition to their high cloning capacity, YACs can be modified by homologous recombination in yeast.

Transgenic mice carrying an intact YAC with the entire Hnf3g locus consistently showed an expression pattern in embryonic and adult tissues indistinguishable from that of the endogenous Hnf3g gene. This is not self-evident, since the heterologous sequences cloned into the Hnf3g locus [β-galactosidase coding region and SV40 intron/poly(A)] might have exerted an effect on transcript processing or mRNA stability. More lines will be needed to further substantiate the observed correct expression from the YAC transgene. This is the first report on position-independent and copy number-dependent expression of a YAC transgene modified with a heterologous reporter gene, which is especially interesting in the light of findings from the human β-globin locus. A locus control region (LCR) located upstream of the gene cluster confers high levels of position-independent, copy number-dependent expression onto linked globin transgenes (for a review see Dillon and Grosveld, 1993). However, when the same LCR was used to drive a β-galactosidase reporter, the LCR lost this property (Robertson et al., 1995; Guy et al., 1996).

Identification and characterization of sequences for tissue-specific expression of Hnf3g

Using the expression data from the Hnf3g YAC as reference, we wanted to identify the sequence elements that are sufficient to produce this pattern. We therefore deter-
Fig. 8. Transient transfections identify a strong enhancer in the 3'-flanking region of the *Hnf3g* gene. (A) A 1.3 kb fragment located at +16 kb displays strong activity in FTO-2B hepatoma cells but not in Ltk- fibroblasts when cloned in front of a *Tk* promoter driving a *Cat* reporter gene. Fragments are aligned with their genomic context (above). For comparison, the deletion introduced in the mutated *Hnf3g* allele (Kaestner *et al.*, submitted for publication) is indicated by a bracket. (B) Definition of a minimal 397 bp *Alu* fragment possessing full enhancer activity in FTO-2B hepatoma cells. The BamHI sites in (B) refer to the 1.3 kb BamHI fragment in (A). *Cat* activities in (A) and (B) were measured relative to pBLCAT5 (Boshart *et al.*, 1992), which was arbitrarily set to 1. Bars represent the mean ± SE from three independent experiments. (C) Sequence of the minimal 397 bp enhancer fragment from (B) with putative binding sites for transcription factors, shown in bold. Restriction sites for a few enzymes are indicated. A, *Alu*; B, *BamHI*; Bs, *Bsu36I*; D, *DpnI*; H, *HindIII*; N, *NheI*; P, *PstI*; X, *XbaI*; Xh, *XhoI*.
Gene regulation of Hnf3g

mined the location of HS in the Hnf3g locus. We tested the functional significance of these sequences by their ability to drive a β-galactosidase reporter gene in transgenic mice. Only construct pγ8Z4.9, containing 3'-flanking sequences (corresponding to the HS at +12 and +16 kb), recapitulated many aspects of the Hnf3g pattern, including expression in liver, pancreas, gut and bone. We were able to dissect the region into two distinct enhancer activities. The functional significance of these sequences by their ability to drive a β-galactosidase reporter gene in trans- Construct pγ8Z3.6, lacking the /H11001 16 kb HS sequence, retained only expression in bone but entirely lost endogenous mice. Only construct pγ8Z4.9, containing 3'/H11032 -flanking sequences (corresponding to the HS at /H11001 12 and /H11001 16 kb), dermal β-galactosidase expression. Interestingly, this was accompanied by appearance of strong ectopic expression recapitulated many aspects of the Hnf3g pattern, including in all transgenic lines, which was not observed in the pγ8Z4.9 construct containing additionally the enhancer active in liver, gut and pancreas. This indicates that the +16 kb region is capable of suppressing position effects, possibly by preventing promoter interactions with neighbouring sequences from the genomic integration site. It has already been described for other genes that deletion or mutation of essential regulatory elements leads to variable transgene expression (Adolph et al., 1993; Bonifer et al., 1994; Ess et al., 1995; Millonig et al., 1995).

The importance of the 3'-flanking region for expression of the Hnf3g gene has been unambiguously demonstrated by examination of the expression of a mutated allele (Hnf3g–) created by homologous recombination in ES cells (Kaestner et al., unpublished results). A deletion encompassing exon 2 and the +12 and +16 kb HS sequences led to silencing of the targeted allele in liver, pancreas, stomach and small intestine. An interesting question arising from this result is whether the Hnf3g–lacZ allele is inactivated in cis or in trans; in other words, is the lack of transcription from the Hnf3g–lacZ allele in these tissues due to lack of the HNF-3γ protein itself. This question can be answered by examining the results obtained from the heterozygous animals (Figure 7A). The bands obtained for Hnf3g arising from the wild-type allele in the heterozygous animals are approximately half as strong as the bands in the wild-type controls, indicating that the reduced level of HNF-3γ protein in the heterozygotes does not impair transcription from the wild-type allele. However, in the same samples, no transcripts originating from the Hnf3g–lacZ allele can be found. Therefore, we conclude that the inactivity of the mutated allele is caused by deletion of cis-regulatory sequences in the 3'-flanking region. Since colon and nasal epithelium staining is observed in both the mutated allele (Hnf3g–) and in pγ8Z4.9 transgenics, we could not yet assign enhancers governing Hnf3g expression in these tissues.

With the results presented above we can ascribe the enhancer activity for expression in liver, pancreas, stomach and small intestine to the +16 kb region, which thus behaves as a posterior foregut and midgut endoderm enhancer of Hnf3g, whereas the +12 kb enhancer is crucial for bone expression.

Fig. 9. Identification of HNF-1 as a strong transactivator of the +16 kb enhancer of Hnf3g. A mutation of the HNF-1 binding site in the context of a 425 bp enhancer construct (in front of a Tk promoter / Cat reporter) (A) causes a dramatic reduction in enhancer activity in transient transfections into FTO-2B hepatoma cells (B). The HNF-1 mutation is indicated by crossed lines. (C) Both HNF-1α (upper) and HNF-1β (lower) activate the 425 bp fragment of the +16 kb enhancer (1 μg 425 bp WT-Cat) but do not show any effect on the fragment with the mutated HNF-1 binding site (1 μg 425 bp mut-Cat) in C33 cervical carcinoma cells. A construct without insert (RSV-O) served as negative control. Cat activities in (B) and (C) were measured relative to pBLCAT5 (Boshart et al., 1992), which was arbitrarily set to 1. Bars represent the mean ± SE from three independent experiments. Numbers in (C) represent amounts (in μg) of the indicated constructs transfected.
Control of Hnf3g expression by HNF-1 constitutes an important component of a transcriptional hierarchy

To define the relevant sequences for the identified posterior foregut and midgut endoderm enhancer activity at +16 kb we have used transfection assays. A minimal 397 bp fragment conferring full enhancer activity in FTO-2B hepatoma cells was identified as containing a functional binding site for HNF-1. We have shown that HNF-1 is a potent upstream activator of the +16 kb enhancer of Hnf3g. Since this enhancer is absolutely crucial for early embryonic expression of Hnf3g in the posterior foregut and midgut endoderm (see above), we wanted to know whether HNF-1 could be involved in activation of the Hnf3g gene. We compared the early expression patterns of the three genes by in situ hybridizations of whole mount and sectioned embryos. In support of earlier reports, we could not detect any Hnf1α expression at the onset of liver formation. Thus, this gene might only have a function in later development, possibly in sustaining Hnf3g expression. The Hnf1β gene, in contrast, is expressed earlier than Hnf3g in the gut endoderm and liver primordium (Ott et al., 1991; Monaghan et al., 1993; data not shown) and would, therefore, be a likely candidate for a direct activator of Hnf3g. Thus, we would like to extend previous models for transcriptional hierarchies in mammalian liver and gut formation (Ang et al., 1993; recently reviewed in Cereghini, 1996; Zarei, 1996). In this model, HNF-3β and HNF-3α are at the top of the cascade and are activated at the onset of definitive endoderm formation. Subsequently, HNF-4 and HNF-1β are induced at the onset of liver differentiation (Ott et al., 1991; Duncan et al., 1994; Taraviras et al., 1994). Hnf3g is then activated by HNF-1β (and possibly other factors). These expression data suggest that a hierarchical relationship exists, however, it cannot be excluded that the cross-regulatory mechanisms described for liver-specific gene expression in adults (for a review see Cereghini, 1996) come into play very early during development.

Materials and methods

Isolation, characterization and modification of an Hnf3g YAC

A 250 kb YAC library from C57Bl/6 mouse DNA (kindly provided by Dr S.Tilghman; Russi et al., 1992) was screened by a polymerase chain reaction (PCR) approach using primers specific for Hnf3g exon 2: γ1, 5′-TTCCAAGGCTTGGAAGTGTTGCA-3′ and γ3, 5′-GTGGCA-GCTTAGTGTTGCA-3′. This resulted in isolation of three independent Hnf3g clones, two of which (designated Yγ4 and Yγ5) had YAC inserts of 170 kb and were characterized in detail. PFGE of several colonies from both clones indicated that they contained stable YACs. In order to verify that the YAC does not contain chimeric sequences, the insert ends were cloned using the ‘vectorette’ PCR method (Riley et al., 1990) and sequenced. Primer pairs derived from both ends were used as sequence tagged sites (STS) to determine their underlying chromosomal origin. Both STS mapped to mouse chromosome 7 (data not shown), where Hnf3g had been localized previously (Avraham et al., 1991; 1994). As a first step towards its modification, the YAC Yγ5 was transferred from the library’s host strain AB1380 to YPH925 by Karcross (Hugert et al., 1994; Spencer et al., 1994), so as to have a broader set of selectable markers. The resulting clone was transformed with the integrating plasmid pRKV1 (Srivastava and Schlesinger, 1991) to destroy the URA3 gene in the right YAC vector arm and replace it with a LYS2 selection cassette. Therefore, URA3 could be used for selection of correct lacZ targeting to the Hnf3g locus by a two-step transplacement strategy (Winston et al., 1983). In the lacZ targeting construct a 6.0 kb Ncol genomic Hnf3g fragment comprising intron and exon 2 sequences was fused in-frame to the lacZ coding region followed by SV40 intron and poly(A) signal (pZaraPC©, kindly provided by Dr L.Montalouli), to obtain pYZ. A 2.6 kb Ncol-Xhol fragment spanning the 3′-untranslated region and flanking sequences of Hnf3g served as 3′-flanking homology. Thus, exon 2 sequences of Hnf3g (position 505–1326 of the cDNA; Kaestner et al., 1994) will be deleted upon homologous recombination with the targeting construct. For transformation, the targeting plasmid was linearized at a unique HindIII site in the 3′-flanking homology region. Transformants were grown on selective medium (Ura-, Trp-, Lys-) and analysed by PFGE, Southern blotting and PCR for the correct targeting event (pop in step in Figure 1B, data not shown). The resulting intermediate YAC harbours duplications of the homology regions that themselves facilitate excision of the intervening sequences when applying selection against the presence of URA3 (pop out step in Figure 1B) using 5-fluoroorotic acid (Sigma). This yields either reversion to the wild-type locus or replacement of exon 2 by the desired exon 2–lacZ fusion. The resultant clone Yγ5Z was verified to have a correct Hnf3g–lacZ fusion without any detectable rearrangements (data not shown).

Generation of transgenic mice

YAC DNA was purified for microinjection from preparative PFGE according to Schedl et al. (1993) with the following modifications. High concentration agarose blocks were prepared as described in Huxley et al. (1991). PFGE conditions were 1% agarose, 0.5× TBE, 14°C, 6 V/cm, using a time ramp with t1 = 15 s and t2 = 20 s and a 20 h running time (BioRad CHEF-DRIII system).

Plasmid insertions were excised from the vector backbone and purified as described (Hogan et al., 1994). Microinjection into pronuclei of fertilized oocytes of FVB/N mice was performed using standard techniques (Hogan et al., 1994).

RNA analysis

Total RNA from adult tissues was isolated after guanidinium isothio cyanate extraction as described (Chomczynski and Sacchi, 1987) RNA concentration was determined by measuring absorbance at 260 nm. RNase protection analysis with 50 pg total RNA was according to Kaestner et al. (1994). The probe used for Hnf3g detection has been described previously (Kaestner et al., 1994). Probes for lacZ gene were generated as follows. A fragment containing the fusion between the Hnf3g and lacZ genes was amplified from DNA obtained from Hnf3g–/– mice using the PCR primers Hnf3g 5′ (TCCCAAGCCTTGGGAC- TGGTCGCA) and lacZ (GGCCATCATGATATGCGC) and sub cloned into the Smal site of Bluescript II (Strategene). Hybridization of an antisense probe derived from this plasmid yields protected fragments of 326 nt from the mutated allele and 65 nt from the wild-type allele.

β-Galactosidase detection

Embryos were dissected in ice-cold phosphate-buffered saline (PBS) and fixed in 0.2% glutaraldehyde, 2 mM MgCl2, 0.02% NP40 in PBS for 1 (E10.5) or 4 h (E12.5) respectively. Embryos were genotyped by PCR from extra-embryonic membranes with primers detecting the lacZ transgene. After fixation, embryos were washed three times for 10 min in PBS and incubated for 1–2 days in staining solution (5 mM K3(Fe(CN))6, 5 mM K4(Fe(CN))6, 2 mM MgCl2, 0.02% NP40, 0.01% sodium deoxycholate, 1 mg/ml Blue-Gal (Gibco BRL) in PBS). Subsequently, embryos were washed three times for 10 min in PBS and post-fixed in 4% paraformaldehyde, pH 7.2, overnight at 4°C. Embryos were washed in PBS, dehydrated and photographed. β-Galactosidase detection on cryostat-sectioned embryos (E14.5) was performed as described in Kaestner et al. (unpublished results).

DNAse I hypersensitivity analysis

Adult mice were perfused with ice-cold PBS and nuclei from liver and kidney prepared according to Becker et al. (1984) and Jantzen et al. (1987). Aliquots of 1.5×107 nuclei were digested with increasing amounts of DNAse I as described (Jantzen et al., 1987). Purified DNA was digested with HindIII, separated on an agarose gel, transferred to a nylon filter and hybridized with a 0.45 kb Xbal–HindIII fragment that indirectly end-labelled a 10.5 kb HindIII fragment in the 3′-flanking region of Hnf3g. Endogenous Xbal, XhoI, BamHI and EcoRI restriction sites from the 3′-flanking region were used to map the identified HS to regions of ~0.5–1.0 kb (not shown). For analysis of the promoter region, genomic DNA was digested with BamHI. A 15 kb BamHI fragment was indirectly end-labelled from its 3′-end located in the intron with a 0.3 kb Xhol–BamHI fragment. HS positions were mapped by comparison with endogenous fragment lengths generated by HindIII/BamHI, BglII/BamHI and NotI/BamHI digestions to regions of ~0.3 kb (not shown).
Plasmid transgenes
A 8.5 kb EcoRI fragment was partially digested with Bsu36I and cloned as an 8 kb promoter fragment fused at position +33 bp (relative to the transcription start site; Kaestner et al., 1994) into the Smal site of pZva+ (Montoliu et al., 1995) to obtain pZF8. For pZF236, a 3.6 kb EcoRI-BamHI fragment from the 3′-flanking region of Hnf3g was inserted into pZva+ of pZF8. Subsequent introduction of a 1.3 kb BamHI fragment extended the cloned 3′-flanking region to 4.9 kb (pZF249). Cloning of an 8.5 kb EcoRI fragment comprising 8 kb promoter, exon 1 and intronic sequences into the HindIII site of pZva (see above) resulted in pZF2Z.

Cell culture and transfections
Rat hepatoma FTO-2B cells (Killary and Fournier, 1984; Killary et al., 1984) were grown in a 1:1 (v/v) dilution of Dulbecco’s modified Eagle’s medium (DMEM) and Ham F12 medium. Ltk– mouse fibroblasts (obtained from ATCC) and C33 human cervical carcinoma cells (ATCC) were cultured in DMEM. Media were supplemented with 10% fetal calf serum, 100 U/ml penicillin, 100 μg/ml streptomycin, 10 mM HEPES, pH 7.4, and 2 mM glutamine. Prior to electroporation, cells were resuspended to 2.2×10^6 cells/ml and mixed with 10 μg plasmid DNA. A Rous sarcoma virus luciferase construct served as internal control. Electroporations were performed as described (Boshart et al., 1990). C33 cells were transfected using the protocol of Chen and Okayama (1987). Cells were harvested after 48 h and Cat and luciferase assays were performed from prepared extracts as outlined in Neumann et al. (1997) and Boshart et al. (1990), respectively.

Accession number
The novel nucleic acid sequence data referred to in this work have been deposited in the DDBJ/EMBL/GenBank Data library under the accession number Y12559.

Acknowledgements
The authors are grateful to Drs A.P.MONAHON, J.A.BLENDY, F.TRONCHE and B.LUTZ for critical reading of the manuscript, Dr F.TRONCHE for help with sequence analysis and Drs L.MONTOLIU and F.TRONCHE for providing plasmids. We would like to thank S.HINZ, S.RIDDER, H.KERN and W.FLEISCHER for expert technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft through grant SFB 229, the Fonds der Chemischen Industrie, BMBF project 0310681 and by European Community grant BI02-CT93-0319.

References


Received on January 29, 1997; revised on April 7, 1997