Clustered arrangement of winged helix genes *fkh-6* and *MFH-1*: possible implications for mesoderm development

Klaus H. Kaestner\(^1, *\), Susanne C. Bleckmann\(^1\), A. Paula Monaghan\(^1\), Johannes Schlöndorf\(^1\),
Antoaneta Mincheva\(^2\), Peter Lichter\(^2\) and Günther Schütz\(^1\)

\(^1\)Division of Molecular Biology of the Cell I and \(^2\)Division of Organisation of Complex Genomes, German Cancer Research Centre, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

*Author for correspondence

SUMMARY

The ‘winged helix’ or ‘forkhead’ transcription factor gene family is defined by a common 100 amino acid DNA binding domain which is a variant of the helix-turn-helix motif. Here we describe the structure and expression of the mouse *fkh-6* and *MFH-1* genes. Both genes are expressed in embryonic mesoderm from the headfold stage onward. Transcripts for both genes are localised mainly to mesenchymal tissues, *fkh-6* mRNA is enriched in the mesenchyme of the gut, lung, tongue and head, whereas *MFH-1* is expressed in somitic mesoderm, in the endocardium and blood vessels as well as the condensing mesenchyme of the bones and kidney and in head mesenchyme. Both genes are located within a 10 kb region on mouse chromosome 8 at 5.26±2.56 cM telomeric to Actsk1. The close physical linkage of these two winged helix genes is conserved in man, where the two genes map to chromosome 16q22-24. This tandem arrangement suggests the common use of regulatory mechanisms. The *fkh-6/MFH-1* locus maps close to the mouse mutation *amputated*, which is characterised by abnormal development of somitic and facial mesoderm. Based on the expression patterns we suggest that a mutation in *MFH-1*, not *fkh-6* is the possible cause for the *amputated* phenotype.

Key words: transcription factor, winged helix, intestinal mesenchyme, somitic mesoderm, *amputated*, mouse, *fkh-6*, *MFH-1*

INTRODUCTION

The 100 amino acid ‘forkhead’ DNA binding domain defines a large and growing gene family with more than 80 members in species ranging from yeast to man (reviewed by Lai et al., 1993). The determination of the crystal structure of the DNA-binding domain of HNF-3\(\gamma\) revealed that DNA recognition is mediated by a variant of the helix-turn-helix motif (termed ‘winged helix’) which contains two loops or ‘wings’ at the C-terminal side of the helix-turn-helix (Clark et al., 1993). The winged helix proteins bind DNA as monomers and make base-specific contacts by the \(\alpha\) helix H3 and loop W2 which bind in the major groove of DNA. A 20 amino acid region N-terminal to helix 3 also has been shown to be important in the determination of binding site specificity (Overdier et al., 1994). Analyses of expression patterns as well as gain- or loss-of-function mutations have implicated the winged helix genes in pattern formation during embryogenesis. The *Hnf-3\(\beta\)* gene product, for example, is normally found in the node, notochord, endoderm and central nervous system during early stages of mouse development (Ang et al., 1993; Monaghan et al., 1993; Sasaki and Hogan, 1993). Ectopic expression of *Hnf-3\(\beta\)* in the midbrain and hindbrain of transgenic mice leads to changes in the expression of floorplate-specific genes and to abnormal neural patterning (Sasaki and Hogan, 1994). Loss of function of the mouse *Hnf-3\(\beta\)* gene through targeted mutagenesis results in severe defects in midline development, specifically the absence of the notochord and floorplate (Ang and Rossant, 1994; Weinstein et al., 1994).

Several other members of the winged helix gene family have been implicated by their expression patterns to play a role in defining mesoderm. Among these are mesoderm/mesenchyme fork head 1 (*MF-1*, Sasaki and Hogan, 1993; also termed *fkh-1*, Kaestner et al., 1993), *MF-2* (Sasaki and Hogan, 1993) and *MFH-1* (Miura et al., 1993). *MF-1* expression was detected in headfold mouse embryos in paraxial mesoderm and later in the mesenchyme of the first branchial arch and declines after day 9.5 p.c. (Sasaki and Hogan, 1993). *MFH-1* gene expression was detected from day 9.5 onward in somites and subsequently in cartilaginous tissues and the metanephros (Miura et al., 1993).

We have previously isolated 16 members of the winged helix family from mice (*Hnf-3 \(\alpha\), \(\beta\) and \(\gamma\)* Kaestner et al., 1994; and *fkh-1* to *fkh-6*, Kaestner et al., 1993, 1995, 1996; Tanoue, Kaestner and Schütz, unpublished data), which exhibit specific and diverse patterns of expression in adult tissues. Here we report the detailed characterisation of *fkh-6*, which was found to be closely linked to *MFH-1* and which together with *MFH-1* defines domains of non-notochordal mesoderm.

MATERIALS AND METHODS

Library construction and cDNA and genomic cloning
An oligo(dT)-primed cDNA library from mouse day 8.5 p.c. (post...
(coitum) embryos (Kaestner et al., 1995) was screened with probes B and D (see Fig. 2) derived from the MFH-1 and fkh-6 containing genomic clone described previously (Kaestner et al., 1993) using high stringency hybridisation and washing conditions (Church and Gilbert, 1984). Five hybridising lambda phages corresponding to MFH-1 were purified and the cDNAs subcloned into Bluescript (Stratagene) and sequenced (Sanger et al., 1977) on both strands. A human placenta genomic library (Stratagene) was screened using probe C (see Fig. 2) to obtain the human MFH-1/fkh-6 locus. Fragments of the genomic phage clone hybridising to fkh-6 and MFH-1 specific probes were subcloned and partially sequenced to confirm the identity of the human fkh-6 and MFH-1 genes.
Fig. 3. Chromosomal localisation of the mouse and human \( fkh-6/MFH-1 \) loci. (A) The murine \( fkh-6/MFH-1 \) locus was mapped to chromosome 8 by linkage analysis as described in Materials and Methods. The numbers refer to the recombinants observed between the \( fkh-6/MFH-1 \) probe and D8Mit35 and Actsk1, as well as the distance in centimorgans (cM) ± the standard error. (B) As revealed by fluorescence in situ hybridisation, the human \( fkh-6/MFH-1 \) locus maps to chromosome 16q22-24. Detection of the hybridised probe via rhodamine revealed highly specific signals as indicated by the arrows in the large panel showing a complete metaphase spread. Panels on the right present additional chromosome 16 homologues from other metaphase spreads to illustrate the subchromosomal localisation on chromosome 16q22-24.

Fig. 4. Expression of the \( fkh-6 \) and \( MFH-1 \) mRNAs during mouse embryogenesis and in adult mouse tissues. 40 µg of total RNA from the adult tissues or whole embryos of the developmental stages indicated were analysed by RNase protection as described in Materials and Methods. The signals obtained were quantified using a phosphoimager and corrected for the specific activity of the probes used and reflect the true molar ratios of the two mRNAs. Note the different scales for the two mRNAs: open bars, \( MFH-1 \) mRNA, scale bar on the left; black bars, \( fkh-6 \) mRNA, scale bar on the right.

Fig. 5. Expression of the \( fkh-6 \) and \( MFH-1 \) genes in day 7.5 (A) and 8 p.c. (B-D) mouse embryos. Whole-mount in situ hybridisation with antisense probes specific for \( MFH-1 \) (A-C) and \( fkh-6 \) (D) was performed as described in Materials and Methods. \( MFH-1 \) is expressed in mesoderm anterior and posterior of the node in late streak embryos (A), and subsequently in presomitic, somitic and head mesoderm (B-C). Transcripts produced by the \( fkh-6 \) gene are detected in day 8.5 embryos in posterior mesoderm (D). Scale bars represent 50 µm in A and 100 µm in B-D. N, node; HM, head mesoderm; H, heart; LP, lateral plate mesoderm, PS, presomitic mesoderm.
**RNA isolation and RNase protection analysis**

Total RNA from a variety of mouse tissues or whole mouse embryos was isolated after homogenisation in guanidinium thiocyanate (Chomczynski and Sacchi, 1987). The quality of the RNA preparations was controlled by ethidium bromide staining of the 18S and 28S rRNAs after electrophoretic separation of the RNA in denaturing agarose gels. RNase protection analysis was performed as described previously (Kaestner et al., 1989) using [α-35S]UTP-labelled antisense RNA probes derived from Bluescript (Stratagene) subclones containing a 556 bp NcoI/BamHI fragment of the fkh-6 gene and a 447 bp Ncol/EcoRI fragment of the MFH-1 gene (probes A and C in Fig. 2). A 337 bp subclone of the mouse TBP (TATA-box binding protein) gene (Tamura et al., 1991) was used as an internal control, as the mRNA of this gene is expressed in all tissues at roughly equal levels, with the exception of ovary and testis. The antisense probes were hybridised overnight against 40 µg of total RNA at 54°C in 80% formamide. Excess probes were removed by digestion with RNases A and T1 and the protected probe fragments analysed on denaturing 6% polyacrylamide gels. The signals obtained were quantified on a Molecular Dynamics phosphoimager and normalised for the number of UMP residues incorporated in the hybridising portion of each probe and represent the molar ratios of the two mRNAs.

**In situ hybridisation**

Mouse embryos and foetuses were obtained from matings between NMRI mice. By convention the day of the appearance of the vaginal plug was counted as day 0.5. Embryos were fixed in 4% paraformaldehyde (pH 7.2) overnight, dehydrated through an ethanol series, cleared in toluene and embedded in paraffin. 6

**Characterisation of the fkh-6 and MFH-1 genes**

In order to understand the relationship of fkh-6 to the other winged helix genes, we determined the primary sequence of the fkh-6 coding region. Although we were not able to obtain full length cDNAs from library screens, we have ascertained that the fkh-6 gene is transcribed and contains only one exon (Figs 3, 4, 5 and data not shown). The fkh-6 mRNA encodes a protein of 337 amino acids with a calculated relative molecular mass of 35870, which is in good agreement with the size of the in vitro transcribed and translated protein (data not shown). The complete sequence is depicted in Fig. 1. The presumed winged helix DNA binding domain is underlined. No striking homology to the transcriptional activation domains identified in HNF-3β (Pani et al., 1992) are present in this sequence, but a large proportion of proline residues (14.8%) was found. A similar proline-rich region was shown to function as transactivation domain in the winged helix protein ‘myocyte nuclear factor’ (Bassel-Duby et al., 1994). The winged helix domain of fkh-6 is most closely related to those of the mouse genes fkh-1 and MFH-1 (77% identity with both genes at the amino acid level; Kaestner et al., 1993; Miura et al., 1993). Outside the DNA binding domain no significant homologies with winged helix or other genes were found in the databases.

During the analysis of the mouse fkh-6 gene we discovered a second winged helix domain located on the same genomic λ phage clone. Sequence analysis revealed that this second winged helix domain corresponds to the previously described MFH-1 cDNA (Miura et al., 1993). Several MFH-1 cDNAs were isolated from a day 8.5 p.c. mouse embryo library which extend 3' to include a consensus polyadenylation signal and poly(A) tail. The complete sequence is depicted in Fig. 2. The start site of transcription was determined by RNase protection assay (data not shown) using probe E (Fig. 2A) derived from the genomic MFH-1 locus. The complete cDNA sequence for the mouse MFH-1 gene encompasses 2340 nt, which together with a poly(A) tail is in good agreement with the size (2.5 kb) of the message estimated from northern blot analysis (data not shown). Sequencing revealed an additional guanine nucleotide at position 49 of the previously published MFH-1 sequence which leads to a new initiator ATG further upstream. The MFH-1 cDNA thus encodes a protein of 495 instead of 462 amino acids and a calculated relative molecular mass of 53081, which agrees with the size (2.5 kb) of the message estimated from northern blot analysis (data not shown). The complete sequence is depicted in Fig. 2. The MFH-1 gene lies approximately 8 kb 5' to fkh-6 and both genes are transcribed in the same direction. Each gene is encoded by only one exon. A similar tandem arrangement has been described for the human winged helix genes HBF-1 and HBF-2, which are both expressed in the brain and which are localised on human chromosome 14 (Wiese et al., 1995) while the other rodent winged helix genes have been found in multiple chromosomal locations (Avraham et al., 1992, 1995; Kaestner et al., 1995). To investigate whether the close linkage of the fkh-6 and MFH-1 genes is unique to mice, we attempted to clone the orthologous locus from a human genomic library and characterised it through hybridisation and sequence analysis. The regions of the human fkh-6 and MFH-1 genes sequenced show 85 and 94% identity at the nucleotide level, respectively, to the corresponding mouse genes. As is shown in Fig. 2, the human fkh-6 and MFH-1 genes are
Expression analysis of the fkh-6 and MFH-1 genes

In order to gain insight into the potential function of the fkh-6 and MFH-1 genes and to assess whether their close linkage results in common usage of regulatory elements governing their expression, we analysed the transcript distribution in adult mice and whole embryos. RNAs from fifteen adult mouse tissues were analysed for expression by RNase protection with probes A an C outlined in Fig. 2. The fkh-6 mRNA is expressed in stomach, intestine and weakly in kidney (Fig. 4). MFH-1 expression is strongest in adult kidney in concordance with the findings of Miura et al. (1993), but is also detectable at lower levels in brain, heart, muscle, uterus, lung, liver, ovary and testis. During embryonic development, both mRNAs are present from day 10.5 onward with decreasing expression after day 11.5 (Fig. 4). The expression of MFH-1 in whole embryos is always 20- to 40-fold stronger than that of fkh-6 (Fig. 4).

In order to precisely define the cellular localisation of the fkh-6 and MFH-1 mRNA in midgestation embryos, in situ hybridisation studies were performed on embryos starting at day 7.5 p.c. of gestation with antisense probes (labelled B and D in Fig. 2). Transcription of the fkh-6 gene is activated in embryos of day 8 p.c. and is found in posterior mesoderm by whole-mount in situ hybridisation (Fig. 5D). Additional expression domains were characterised by in situ hybridisation of sectioned embryos. Transcripts were not detectable in early headfold embryos (Fig. 6A), but appear on day 9.5 of gestation in the mesenchyme surrounding the anterior gut and lung as well as mesenchyme of the palate (Fig. 6G). This expression is maintained until day 12.5 p.c. (see below).

Whole-mount in situ hybridisation demonstrates that the MFH-1 gene is activated prior to fkh-6 in late primitive streak stage embryos in the mesoderm both anterior and posterior to the node, but not in the node itself (Fig. 5A). The signal extends to the primitive streak but is excluded from the allantois and extraembryonic regions. In day 8 p.c. embryos, just prior to turning, expression is weak in the head mesenchyme and becomes stronger in somitic and presomitic mesoderm (Fig. 5B). After turning, expression is maintained in the somites and increases in head mesenchyme (Fig. 5C). In situ hybridisation of sectioned embryos demonstrates that MFH-1 expression is absent from day 6.5 embryos (data not shown), but is present in the mesoderm of late primitive streak and headfold stages (Fig. 6C,D). No transcripts are detectable in the endoderm, surface ectoderm or in the extraembryonic regions of the embryo. In 8 to 8.5 day old embryos, MFH-1 mRNA is localised to all mesenchymal lineages but subsequently becomes localised to the somites, presomitic mesoderm, head and endocardial mesoderm (Fig. 6E). At later stages, MFH-1 transcription is maintained in the mesenchyme from head to tail, including the dorsal aorta (Fig. 6H) and the somite mesoderm, where it becomes localised to the sclerotome (Fig. 6H and data not shown), but is excluded from gut mesenchyme.

In order to allow the assessment of potential overlap in the expression of the two genes, we performed in situ hybridisation studies on adjacent sections of day 12.5 p.c. mouse embryos with antisense probes for MFH-1 and fkh-6. As is demonstrated in Fig. 7, transcripts for both genes are found in partially overlapping regions of the head mesenchyme, for instance in the precartilage primordia of the temporal bones, the nasal, otic and optic capsule, the nasal septum and the tongue (Fig. 7A,B,E,F). Transcripts of the MFH-1 gene are, however, more abundant than those of fkh-6 and are additionally localised to the mesenchyme surrounding the trachea and the cartilage primordia surrounding the neural tube (Fig. 7F), while fkh-6 mRNA is exclusive to the teeth. In the trunk, fkh-6 transcripts are found in the mesenchyme surrounding the notochord, the mesenchyme of the lung and developing kidney, the wall of the intestine and stomach and at a low level in the condensing cartilage of the bones (Fig. 7C,D). The MFH-1 gene is transcribed in the condensing cartilage of the tail and limb bones, the blood vessels and the urogenital system (Fig. 7G,H).

In summary, there are overlapping expression domains of the two genes of the fkh-6/MFH-1 locus in the mesenchyme of the head, bones and urogenital system, but there are also areas of mutual exclusion. For example, MFH-1 is a marker for developing blood vessels, while fkh-6 is specific for the mesodermal components of the gut.

**DISCUSSION**

**Is the fkh-6/MFH-1 locus orthologous to sloppy paired of Drosophila?**

Characterisation of cDNA and genomic clones of the mouse...
fkh-6 and MFH-1 genes have localised these genes in a close tandem array, which suggests common regulatory mechanisms or functional redundancy. A similar clustering has been described for the winged helix genes sloppy paired 1 and 2 of Drosophila (slp1 and slp2; Grossniklaus et al., 1992). Mutation of the sloppy paired locus leads to a phenotype reminiscent of the pair-rule class of segmentation genes with impaired involution of the head and partial deletion of naked cuticle in the odd-numbered abdominal segments and the mesothorax. The expression pattern of the two genes is complex but very similar, with the activation of slp1 preceding that of slp2. In these studies it was demonstrated that until gastrulation only slp1 is required, while at later stages the two genes are redundant in many respects. Based on these observations the authors concluded that the two slp proteins are biochemically equivalent and that the greater requirement for slp1 is explained in large part by its earlier expression (Cadigan et al., 1994).

Could a similar relationship exist between MFH-1 and fkh-6? The MFH-1/fkh-6 locus does not appear to be orthologous to the sloppy paired locus despite the similar structure of the loci. The sequence similarities between the Drosophila and mouse genes are limited (only 62 or 58% identical amino acids in the winged helix domain between MFH-1 and slp1 and slp2, respectively, versus 72% amino acid identity between MFH-1 and Drosophila forkhead; Weigel et al., 1989) and the characteristic carboxy-terminal sequence of the slp genes is not present in MFH-1 or fkh-6. In addition, the expression domains...
are dissimilar. Furthermore, in contrast to the *slp* genes, the two mouse genes have only partially overlapping expression patterns and are probably not redundant (see below).

**The *fkh-6* and *MFH-1* genes define overlapping mesodermal domains**

The *fkh-6/MFH-1* locus contains two genes which are expressed in partially overlapping domains. Transcripts for both genes are found in the mesenchyme of the developing head, kidney and bones. The *fkh-6* gene is active in addition in the mesenchymal layer of the gut, but is excluded from the gut epithelium, which is derived from the definite endoderm. Proper differentiation of the intestinal epithelium is thought to depend on the establishment of functional interactions with the adjacent mesenchyme (Birchmeier and Birchmeier, 1993). A possible role of the *fkh-6* gene could lie in the establishment or maintenance of such interactions, or more specifically, in the establishment of an inductive signal from the gut mesenchyme to the epithelium. Preliminary results from our analysis of mice homozygous for a targeted mutation in the *fkh-6* gene suggest that this interaction has been impaired. Mice lacking this gene exhibit abnormal development of the stomach and small intestine, both areas with relatively high and exclusive *fkh-6* expression (Kaestner and Schütz, unpublished data).

The *MFH-1* gene, in addition to its expression domain in the head mesenchyme, seems to play a role in somite formation, as it is expressed in paraxial mesoderm from its inception, and later on in pre-somatic as well as somitic mesoderm, where transcripts become localised to the sclerotomes. In addition,
MFH-1 transcripts are also localised to the developing heart and blood vessels as well as the kidney and mesenchyme of the limbs, the latter expression domains were also described by Miura et al. (1993).

The regulation of the fhk-6/MFH-1 locus must necessarily be complex to accommodate the similar yet divergent expression patterns of the two genes. It seems plausible that a common enhancer element could govern the expression of the two genes in the head mesenchyme, but that additional regulatory elements may be required for the individual expression domains of the two genes.

We would like to thank Heidrun Kern and Evelyn Grau for expert technical assistance, Werner Fleischer for oligonucleotide synthesis, Drs J. A. Blendy and B. Lutz for critical reading of the manuscript and the European Backcross Collaborative Group for the help in the linkage analysis. This work was supported by the Deutsche Forschungsgemeinschaft through SFB 229, the Fonds der Chemischen Industrie and by European Community grant number BI02-CT93-0319.

REFERENCES


(Accepted 4 March 1996)