FHX, a Novel Fork Head Factor with a Dual DNA Binding Specificity*

(Received for publication, August 19, 1999, and in revised form, January 27, 2000)

Cristina Pérez-Sánchez‡‡, María Ana Gómez-Ferrería‡, Carmen Arias de la Fuente‡‡, Begona Granadino‡‡, Gloria Velasco, Andrés Esteban-Gamboa‡‡‡, and Javier Rey-Campos‡‡

From the ‡Centro de Investigaciones Biológicas, CSIC, Velázquez 144, 28006 Madrid, Spain and ‡Departmento de Bioquímica y Biología Molecular, Facultad de Medicina, Universidad de Oviedo, 33006 Oviedo, Spain

The HNF3/fork head family includes a large number of transcription factors that share a structurally related DNA binding domain. Fork head factors have been shown to play important roles both during development and in the adult. We now describe the cloning of a novel mammalian fork head factor that we have named FHX (fork head homologous X (FHX), which is expressed in many adult tissues. In the embryo, FHX expression showed a very early onset during the cleavage stages of preimplantation development. Polymerase chain reaction-assisted site selection experiments showed that FHX bound DNA with a dual sequence specificity. Sites recognized by FHX could be classified into two different types according to their sequences. Binding of FHX to sequences of each type appeared to occur independently. Our data suggest that either different regions of the fork head domain or different molecular forms of this domain could be involved in binding of FHX to each type of site. In transfection assays, FHX was capable of activating transcription from promoters containing FHX sites of either type.

The regulated expression of the genome is essential for the homeostatic maintenance and correct cell differentiation and morphogenesis of an organism. Transcription factors play a preponderant role in this regard. From an evolutionary standpoint, all transcription factors derive from a small set of primitive factors. These, upon successive gene duplication-divergence events and fusion with other genes or parts of genes, have given rise to the whole repertoire of fork head factors. Consequently, current transcription factors may be sorted into different families according to their homology relationships. One of these families, which appears to have been highly successful in evolution, is the fork head/HNF3 family, with known members in species from yeast to humans, with the striking exception of green plants, where no fork head factor has been described so far (for a review, see Ref. 1). Common to all fork head factors is their structurally similar DNA binding domain. This domain, known as fork head or winged helix domain, folds in a particular three-dimensional structure resembling those of the prokaryotic Escherichia coli BirA, CAP, and Ompr factors (2, 3), the eukaryotic Ets (4) and Rap30 DNA binding domains (5), and also the globular part of the chicken histone H5 (6). Despite their similar topologies, these proteins do not display any clear sequence similarities. They seem more related than the result of function-driven evolutionary convergence than a consequence of a common evolutionary past. Drosophila fork head was the first member of this family identified as a gene implicated in the development of gut and head structures of the fly (6). In mammals, the first fork head homologues described were the liver enriched transcription factors HNF3α, β, and γ (7, 8), initially identified as regulators of the expression of several liver-specific genes.

In vitro, fork head factors bind DNA as monomers, and they appear to induce a strong bending of the DNA molecule (2, 9). HNF3β has been shown to be capable of binding to its cognate sequence even when this is packed in nucleosomes (10, 11). This binding results in nucleosome destabilization, which endows HNF3β with a chromatin remodeling function. This nucleosome-destabilizing activity may be related to the DNA binding capacity demonstrated for some of these factors and could be a property of the fork head domain itself.

Fork head factors have been shown to play important roles in cell differentiation and developmental processes (12–15), control of life span (16, 17), circadian rhythms (18), and signal transduction (17, 19–21). However, not much is known about the genes that are the final players of these phenomena, which are regulated by fork head factors. DNA sequences recognized by many fork head factors show a conserved motif with the consensus (T/C)AAACA. Differences in the recognition sequences for different fork head factors usually lie at positions flanking this core element. Nevertheless, there are examples where several fork head factors can bind to the same DNA sequence (9). This apparent promiscuity in DNA binding potentially allows for functional redundancy between some fork head factors. This has been shown to be the case for HNF3γ in the mouse (22). HNF3γ is a knockout mice are fully viable and fertile with an apparent mild phenotype characterized by a 30–50% lower expression of some hepatic genes. It has been proposed that, in this case, the absence of HNF3γ is ameliorated by an augmented expression of the homologous HNF3α and -β. However, in the reciprocal case, the absence of the HNF3β gene could not be substituted by the normal complement of either HNF3γ or -α (23). In many other examples where this issue has been approached, the whole repertoire of fork head factors...
could not rescue the phenotype produced by the absence of one single factor. This suggests that the possible functional redundancies between members of this family most probably restrict to some highly homologous pairs with overlapping expression and regulation.

We have addressed the characterization of human fork head factors potentially involved in the biology of the liver, in the adult or during development. In this study, we describe a novel member of this family, which we have named FHX (fork head homologous X). FHX bound DNA in a specific manner. PCR-assisted site selection experiments showed that this fork head factor was able to bind to two different types of sequences. On the one hand, it bound to sequences that, although specific for FHX, contained the core element also found in sites for some other members of the fork head family. However, a number of FHX-specific sites showed sequences that did not contain this core element and could not be aligned with sequences of the other type. The binding of FHX to sequences of each type showed differences that suggest that this fork head factor could bind DNA in two different, independent, manners. In any case, irrespective of the type of target sequence used, FHX was capable of activating transcription in transfection experiments.

### EXPERIMENTAL PROCEDURES

#### Screening of cDNA Libraries and Sequencing

Human cDNA libraries (adult and fetal liver and adult brain) were screened according to standard procedures (24). High stringency washes (1% SDS at 65°C) were used. Nucleotide ends were determined with an Applied Biosystems 377 automated sequencer. Other general molecular biology techniques were performed following established procedures (24).

#### Expression Constructs and Reporters

**PGK.FHX**—A fragment of the FHX cDNA encompassing just the coding sequence was cloned, under the control of the phosphoglycerate kinase 1 (PGK) promoter (25), between the XhoI and XbaI restriction sites of a PGK promoter plasmid vector.

**RSV.Gal4::FHX**—A fragment of the FHX cDNA, from the initiator ATG to position 1113 of the coding sequence was cloned in frame downstream of the DNA binding domain of the yeast transcription factor Gal4, in an expression vector under the control of the Rous sarcoma virus (RSV) promoter and a TATA box (26). A fragment of the first AUG codon of the FHX coding sequence (UAC-ATATTTTTAACTTGCTT-3') was used as a probe in Northern blotting experiments. The coding region for the protein was amplified by PCR from human fetal liver cDNA, using the primers UAC-ATATTTTTAACTTGCTT-3' and 5'-CCGCTGAGGGCCATGAGG-3' (only one strand shown). Primers used for PCR amplification were 5'-GGCGCTGAGGGCCATGAGG-3' and 5'-GGCGTCGACATTCTAGAAGG-3' (only one strand shown).

#### Expression of Recombinant Fusion Proteins

The FHX fork head domain was expressed in *E. coli* as a fusion protein with GST. A 320-base-pair *Sma*I/*Mun*I restriction fragment of the FHX cDNA coding residues Gly37 to Ile142, was cloned in the *Sma*I and EcoRI restriction sites of the *pGEX-3X* vector (Amersham Pharmacia Biotech) in frame with the GST open reading frame. Recombinant bacteria were grown overnight in Terrific broth (24), in the presence of 150 μg/ml ampicillin and 1% glucose, at 37°C with strong agitation. Cultures were centrifuged, and the bacteria were resuspended in fresh medium containing 150 μg/ml ampicillin and 0.5 mM isopropyl-1-thio-β-D-galactopyranoside and induced for 4–16 h at 37°C with agitation. Bacteria were centrifuged, resuspended in phosphate-buffered saline, and lysed in a French press. Triton X-100 was added at 1%, and the lysate was cleared by centrifugation. GST::FHX fusion protein was recovered from the supernatant by affinity chromatography on glutathione-agarose columns (Sigma). Elution of the protein was performed overnight at 4°C with reduced 50 mM glutathione in 250 mM Tris-HCl, pH 8.0. The purified GST::FHX fusion protein was dialyzed against 50 mM Tris-HCl, pH 8.0, 2 mM dithiothreitol, 15% glycerol, to remove the free glutathione used in the elution step and used for the PCR-based site selection experiments and EMSAs.

#### DNA Binding Assays

FHX binding to DNA was assayed using standard EMSA procedures (24). Briefly, crude bacterial extracts or purified recombinant GST::FHX fusion proteins were incubated with radiolabeled oligonucleotide probes in a buffer containing 10 mM Tris-HCl, pH 7.9, 50 mM NaCl, 2.5 mM MgCl₂, 1 mM EDTA, 1 mM dithiothreitol, and 4% glycerol. 1 μg of poly(dI-dC)/poly(dI-dC) was added as a nonspecific competitor. Double-stranded oligonucleotide probes were labeled by filling the 5' protruding ends with [γ-32P]dCTP using the Klenow fragment of DNA-polymerase I. Between 0.1 and 1 ng of probe was used per assay.

#### PCR-based Site Selection

Selection of DNA sequences specifically recognized by the FHX fork head domain was done following a PCR-based protocol as described (9, 26). We used a pool of double-stranded 59-mer oligonucleotides that contained the central 19 bases fully randomized. The ends of these degenerated oligonucleotides contained fixed known arbitrary sequences, which allowed us to use them for PCR amplification with specific primers. The generic sequence of the starting mix of oligonucleotide probes was 5'-GCCGCTGAGGGCCATGAGG-3' (only one strand shown). Primers used for PCR amplification were 5'-GGCGCTGAGGGCCATGAGG-3' and 5'-GGCGTCGACAGCTTTGCTAGA-3'. Purified GST::FHX fusion protein (30 ng) was incubated with 1 ng of the mixed probe for 15 min on ice in the same buffer used for EMSAs. FHX::GST-bound oligonucleotides were affinity-purified onto glutathione-agarose beads. After extensive washes with phosphate-buffered saline, oligonucleotides were eluted, purified by electrophoresis on 4% low sieve agarose gels, and subsequently used as probes for the next selection cycle. A total of five selection cycles were performed. Oligonucleotides from the final selection step were cloned in the EcoRV site of pBluescript SK- plasmid. The inserts of more than 50 independent clones were sequenced. Sequences were first automatically aligned using the ClustalX program (27). The alignments were further refined manually.

#### Transfections

Cell transfection experiments were performed in the human hepatoma cell line Hep3B. Cells were grown in 1-cm diameter culture wells and transfected either using the LipofectAMINE reagent (Life Technologies, Inc.) or the calcium-phosphate method. A total amount of 1 μg of DNA was used in each transfection experiment. Luciferase activity was determined using the Single Luciferase assay kit from Promega. Transfection efficiency was normalized by co-transfecting a PGK promoter-driven β-galactosidase reporter plasmid.

### RESULTS

#### Isolation of cDNA Clones for FHX

In order to identify novel human fork head factors, we used a fragment of the rat HNF3γ cDNA as a probe to screen a λ-gt10 cDNA library derived from human fetal liver. Several positive clones showing strong hybridization signals were obtained, the inserts of which were PCR-amplified and directly sequenced. As expected, some of these clones corresponded to the human HNF3γ cDNA. However, one clone, H3.1, did not show sequence similarity with HNF3γ, except for the fork head domain. BLAST (28) searches of the GenBank™ with the sequence of H3.1 showed strong similar, but not identity, with transcription factors of the fork head/HNF3 family, suggesting that this clone corresponded to a novel member of this family. We named this novel fork head factor “fork head homologous X.” The 1.8-kilobase partial insert of the H3.1 clone was used to probe the obtain targeted clones that extended differently in both 5' as well as 3' ends. From these overlapping clones, encompassing a total of 4860 nucleotides of unique sequence, the full coding region for FHX was derived, which spanned 1722 nucleotides and coded for 574 amino acid residues (Fig. 1). The surrounding sequence of the first AUG codon of the FHX coding sequence (UAC-
CAUGG) fitted perfectly the Kozak criteria for efficient eukaryotic initiator AUGs (29). No other in frame AUG was found within the 489 nucleotides upstream of this AUG. In addition, an in frame UAG stop codon was located 425 nucleotides upstream, which discarded the possibility of translation initiation from unidentified further upstream AUGs. This strongly suggested that this AUG could be the actual initiator codon of the FHX protein.

Amino acid sequence alignments between the putative fork head domain of FHX and those of other fork head factors revealed many amino acid residues only present in FHX. Still, secondary structure predictions were compatible with the x-ray-derived three-dimensional structure obtained for HNF3γ. Furthermore, computer-assisted molecular modeling of the FHX fork head amino acid sequence fitted very well the structure (root mean square, 0.997 Å) of the HNF3γ fork head domain crystal (2). The most dissimilar region of the FHX fork head sequence with respect to other fork head factors was the region equivalent to the W2 “wing” (Fig. 2A). The corresponding region of Genesis, another fork head factor whose structure has been recently determined by NMR studies (30), has been shown to be highly disordered. This may suggest that this region of the fork head domain could have fewer structural constraints. In addition, we found that for FHX, this region was not necessary for specific DNA binding (see below). Therefore, in this report, we have not included this region as a part of the fork head domain of FHX; this is reflected hereafter and in the figures.

Homology searches performed with the C-terminal part of FHX did not reveal any significant homology with any reported sequence. A conspicuous motif within this part of the molecule was a region of extremely high content in proline and glutamine residues. Although the functional significance of this region is so far unknown, this abundance in prolines and glutamines is reminiscent of transactivation domains of other transcription factors and also of regions involved in protein-protein interactions. The N-terminal part of FHX, upstream to the fork head domain, has no resemblance to any previously characterized structural motif. However, amino acid residues 5–20 of FHX (domain I) shared a striking homology with a region of the N terminus of HFH4 (31), also conserved in fork head factors HTLF (32), CHES-1 (33), and Trident/WIN/HFH11 (34–36) (Fig. 2B). FHX also shared with HFH4 a 7-amino acid region (from position 150 to 156 of FHX and from 212 to 218 of HFH4).

**Fig. 1.** Nucleotide and amino acid sequences of FHX. Sequence of the FHX cDNA and its conceptual translation are shown. Open boxes indicate domains I and II (see “Results” and Fig. 2); the region homologous to the fork head domain is indicated with a filled box. The glutamine- and proline-rich region is boxed in gray. The putative ATG initiator codon is shown in boldface type. The sequence similar to the Kozak consensus for initiator AUGs of eukaryotic mRNAs is underlined. Terminator codons are indicated with filled boxes. Numbers in boldface type to the right correspond to nucleotides of the coding sequence. Italicized numbers indicate the last amino acid of each line. A schematic representation of FHX is shown below.
to 218 of HFH4) rich in basic residues (domain II). Basic regions are also present in many other members, but not all, of the fork head/HNF3 family, and for HNF3β (37) and FREAC-2 (38) they have been proposed to be important for nuclear translocation. HFH4 was also the member of this family that displayed the highest degree of sequence homology with FHX in the fork head domain (Fig. 2A).

FHX mRNA was detected in all human tissues assayed by Northern blot analyses (data not shown). It migrated as a single band of approximately 6.5 kilobase pairs. Reverse transcription-PCR experiments with total RNA from murine embryos showed that FHX expression started very early in pre-implantation development. Thus, although it could not be detected in one-cell embryos, the FHX mRNA could be easily detected in eight-cell stage embryos (not shown).

**Binding of FHX to DNA—** Sequence conservation in the fork head domain strongly suggested that FHX could bind DNA specifically. To determine the DNA sequences recognized by FHX, we performed site selection experiments following a PCR-based strategy (9, 26) as described under “Experimental Procedures.” For these experiments, we used a recombinant fusion protein of FHX with glutathione S-transferase (GST), GST::FHX, which comprised from Gly42 to Ile42 of FHX and thus included the fork head domain but not the region equivalent to the “wing” W2 (Fig. 3A). The enrichment of FHX-specific sites obtained in the selection procedure is illustrated in Fig. 3B. Several individual clones of the selected FHX-specific sites were sequenced. Two different sets of sequences were found. Sequences of the largest set, named type A FHX sites, were aligned, and a consensus was derived (Fig. 3, C and D). This consensus showed a core element, (A/G)/(T/C)AAA(C/T)A, similar to that described for other fork head proteins (9, 24). This consensus showed a core element, (A/G)(T/C)AAA(C/T)A, similar to that described for other fork head proteins (9, 24). Careful analysis of the sequence alignments showed that certain nucleotides 5′ to the core element appeared to have been more enriched than those at the 3′ end, from which a consensus could not be inferred. However, despite this apparent relaxed specificity downstream from the core element, some nucleotide residues were clearly less represented among the sequences studied. For example, at position +9 (first base of the core element is position +1) “G” was found in only 5% of the sequences. A small number (n = 6) of the selected sequences, named type B FHX sites, differed significantly from those discussed above and could not be aligned with confidence (Fig. 4A). Whereas there were obvious similarities among these divergent sequences, their number was too low to derive a definitive consensus. Most noticeably, these sequences did not show the well conserved (A/G)/(T/C)AAA(C/T)A core element. Strikingly, they were still able to bind FHX. DNase I footprinting analyses with the GST::FHX fusion protein and a probe containing one of these infrequent FHX sites, showed a protected region that was not observed in control experiments with the GST protein alone (Fig. 4B).

In order to find target genes putatively regulated by FHX, we used a custom made computer program to search a data base of vertebrate gene-regulatory regions, utilizing as a probe a frequency matrix of each nucleotide at each position of the type A FHX sites. This search resulted in a rather large set of genes that contained putative FHX sites. We wanted to assess the binding capacity of FHX to these naturally occurring sites. To this end, we performed electrophoretic mobility shift assays (EMSA) with a double-stranded oligonucleotide probe containing a putative type A FHX site from the promoter region of the human gene C4BPA, which encodes the α-chain of the complement regulatory protein C4b-binding protein (40). As Fig. 4C shows, incubation of this probe with the GST::FHX fusion protein gave rise to a retarded complex, the formation of which could be competed with a 100-fold excess of the same unlabeled oligonucleotide, but not with oligonucleotides containing binding sites for other transcription factors.

**FHX Bound Type A and Type B Sequences Differently—** Since both type A and type B FHX sites differed considerably in sequence, we wanted to know to which type of sequences FHX bound better. To this end, we performed Scatchard plot analyses of the FHX binding to representative sequences of each FHX site type. As a type A site, we used the FHX site of the C4BPA promoter. The sequence shown in boldface type in Fig. 4A was used as a type B site. We carried out quantitative EMSA experiments with a fixed amount of affinity-purified GST::FHX recombinant fusion protein and variable amounts of each 32P-labeled probe. As shown in Fig. 5A, FHX bound to type A and type B probes in a dose-dependent manner and with comparable, although not identical, relative binding affinities for both types of probes. Of note, however, FHX showed a discrete, but reproducible, higher binding affinity for the type B probe (0.0074 fmol−1) compared with the type A probe (0.0051 fmol−1) (Fig. 5B). This is apparently in contrast with the results obtained in the PCR-assisted site selection experiments, shown in Fig. 4, where the relative frequencies of the type B sequences were lower than those of the type A sites.

As a different approach to answer the question of to which type of sequences FHX bound better, we performed competitive EMSA experiments with labeled probes of each type, using both unlabeled oligonucleotides as competitors. The results of these experiments are shown in Fig. 6. Interestingly, binding of FHX to type A probe was better competed by the same unlabeled oligonucleotide than by a type B oligonucleotide. Reciprocally, binding to type B probe was competed more efficiently by the same type B oligonucleotide than by type A sequences. This suggested that FHX could bind each type of sequence in a different manner (see “Discussion”), and thus competition of FHX binding to one sequence with sequences of the other type would not be as efficient as with the homologous sequence. In
addition, as can be observed in Figs. 5 and 6, depending on the probe type used, the FHX-DNA complexes showed different electrophoretic behaviors. 

Activation of Transcription by FHX—The experiments described above showed that FHX was capable of specifically binding DNA. To determine the capacity of FHX to regulate transcription in a DNA-binding dependent manner, we performed transient transfection experiments in Hep3B human hepatoma cells. We used a Luciferase-based reporter gene construct, with five FHX sites cloned in front of a minimal promoter containing just a TATA box. The FHX sites used in these initial experiments were the type A FHX site of the human C4BPA promoter (we named this construct 5×FHX-A.Luc). 

Co-transfection experiments of this reporter and increasing amounts of an expression vector of FHX showed a dose-dependent increase of the reporter gene expression (Fig. 7B).

The binding of FHX to type B sites also resulted in activation of transcription. Thus, when a similar reporter construct (but with four type B FHX sites cloned in front of the same TATA-box containing minimal promoter (named 4×FHX-B.Luc)) was used in the co-transfection experiments, an increase in the reporter activity was readily observed. Although the FHX-dependent transcription of the 4×FHX-B.Luc reporter construct was lower than in the case of the 5×FHX-A.Luc (Fig. 7C), the basal activity of the type B reporter was also lower, which resulted in a practically identical induction ratio for both re-
porter constructs (Fig. 7D). This suggests that the binding of FHX to either type of the FHX sites has a similar effect on transcriptional activation.

We observed that rather high doses of the FHX expression vector (1:4 reporter/effect ratio) had to be used to clearly see the activation effect of FHX with these synthetic reporter systems. This might be due to the endogenous FHX expressed in the recipient Hep3B cell line, which could contribute to raise the threshold level necessary for exogenous FHX to produce a clear effect on transactivation. In addition, other transcriptional activation by FHX.

A, schematic representations of the reporter and effector constructs used in transfections shown in B–D. Reporter A (5×FHX-A.Luc) had five type A FHX sites cloned in front of a TATA-box-containing minimal promoter controlling the firefly luciferase reporter gene. Reporter B (4×FHX-B.Luc) was similar to reporter A but with four type B FHX sites cloned in front of the TATA-box. B, dose-response analysis of the FHX-mediated transcriptional activation of the reporter A shown in A. Hep3B cells were transfected with 0.2 mg of the 5×FHX-A.Luc reporter plasmid and different amounts of a PGK.FHX expression vector. We used a PGK.Neo plasmid to maintain the total amount of effector DNA at 0.8 mg in each transfection. For each point, the mean and the S.D. (error bars) of at least three independent experiments are shown. The inset shows the exponential dose-response curve of transcriptional activation by FHX. C and D, FHX-mediated transcriptional activation of promoters with type A and type B FHX sites (reporters A and B shown in A). Hep3B cells were transfected with the reporter constructs 5×FHX-A.Luc (reporter A) and 4×FHX-B.Luc (reporter B), which contain the luciferase gene driven by synthetic promoters with five copies of type A and four copies of type B FHX sites, respectively, cloned in tandem upstream of a TATA box and the FHX expression vector PGK.FHX. A reporter/effect ratio of 1:4 was used in these experiments. C, the transcriptional activation observed from each construct in relative Luciferase units. D, the relative induction of the transcription of each reporter construct, by FHX. The mean and the S.D. of at least three experiments are shown. E and F, transactivation of a Gal4-dependent luciferase reporter system by an expression vector of a Gal4DBD::FHX chimeric protein. Hep3B cells were transfected as in B but using the 5×Gal4.Luc reporter plasmid and two different amounts of the effector construct (shown in E). As a control, the same amounts of an equivalent Rous sarcoma virus-null plasmid were used in parallel transfections. As in B, the mean and the S.D. of at least three experiments are shown.

FIG. 5. Scatchard analyses of the FHX binding to type A and type B sites. A, EMSAs performed with decreasing amounts of labeled double-stranded oligonucleotide probes containing a type A (probe A, 5′-GGAAGAGCTTAAACAGTGCTGCTT-3′) or type B (probe B, 5′-GGAGAGACATAGTTTTATTAAACAGTGCTGCTT-3′) FHX site and a constant amount of purified recombinant GST::FHX fusion protein. Arrowheads b and f indicate the FHX-bound and free probe, respectively. The total amount of probe used in each lane (from 1 to 13) was 76, 57, 38, 30.4, 22.8, 15.2, 7.6, 3.8, 3.04, 2.28, 1.52, 0.76, and 0.38 fmol, respectively. Lane 0 corresponds in each case to probe alone (38 fmol). B, Scatchard plot of FHX binding to probes A and B shown in A. Bound and free probe were measured using a PhosphorImager and plotted. Relative affinity constants for FHX binding to probe A and probe B were taken as the slope of the linear regression fit of the data. Only values in the linear part of the curve were taken into account for these calculations.

FIG. 6. Competitive EMSAs with type A and type B probes and purified recombinant GST::FHX fusion protein. Double-stranded oligonucleotides containing type A and type B (probe A and probe B, respectively) were used as probes, and their binding to purified recombinant GST::FHX fusion protein was competed with increasing amounts of the same or the counterpart unlabeled oligonucleotides.
factors present in the cell, which could bind to the FHX sites in the reporter plasmid, would also compete with FHX for binding to the reporter and so could contribute to the low response observed with low concentrations of transfected FHX. In keeping with this hypothesis, when an expression vector of a fusion protein, between FHX and the DBD (DNA binding domain) of the yeast transcription factor Gal4, was co-transfected in Hep3B cells with a Luciferase reporter plasmid with five Gal4 sites, we observed activation of transcription at low doses of the effector plasmid (even at a reporter/effector ratio of 1:0.1. Fig. 7F). In this case, the binding to DNA of the hybrid molecule Gal4DBD::FHX could not be competed out by any other endogenous protein, thus explaining why activation of transcription was readily observed even at low doses.

**DISCUSSION**

Here we report the cloning and characterization of FHX, a novel human fork head transcription factor. Apart from the fork head, the FHX sequence did not show any significant homology to any other gene in the data banks. This led us to conclude that FHX is indeed a novel gene and not the human homologue of another species fork head gene. From amino acid sequence comparisons, FHX fork head is most similar first to the fork head domain of HNF-4 factor (41) and, second, to Trident/WIN/HFH-11 (34–36), CHES1 (33), and HTLF-1 (32) fork head factors.

PCR site selection experiments with recombinant FHX indicated that the binding of this fork head factor to DNA had a dual sequence specificity. Two different types of sequences were obtained, which showed specific binding to FHX. Type A sequences contained a core element also found in sites for other fork head factors. Type B sequences, found in a lower number, did not show this core element and could not be aligned with the type A sites. Although the frequency of type B sequences in the site selection experiments was lower than that of type A sequences, both bound FHX with comparable affinities. Furthermore, for both types of sites, competition of the FHX binding to sequences of each type was always better when the competitor sequence was of the same type than when it was of the other type. This can only be explained if the binding of FHX to both types of sequences occurs through two independent mechanisms. This would happen, for example, if sequences of each type interact with different regions of FHX. Since in the site selection experiments we used just the fork head domain of FHX fused to GST, the specific interactions of FHX with both types of sequences probably do not require regions of the protein other than the DNA binding domain itself. It would still be possible that different parts of the fork head are involved in the interaction with each type of sequence. Another possibility also compatible with our data would be that two different molecular forms of the FHX fork head interact with sequences of each type. If so, each molecular form of FHX could have a higher affinity for one sequence type than for the other. This would explain why the homologous competition is more efficient than the heterologous in the EMSAs and, despite the similar affinities of both sequences for FHX, the frequency of sequences of each type in the site selection experiments could be different, since this would also depend on the relative abundance of each molecular form of the FHX protein in the assay. Also compatible with this interpretation is the fact that the electrophoretic behaviors of the FHX-DNA complexes with each sequence type in the EMSAs are different.

Transfection experiments showed that FHX was capable of activating transcription. The level of FHX-dependent transcription of promoters containing sequences of both types of FHX sites was similar. This suggests that the binding of FHX to type A and type B sites, whatever the differences in the molecular mechanisms involved, permits the activation of transcription mediated by the possible transactivation domains of FHX. However, it is still unknown whether the binding of FHX to each type of sequence may have consequences in the regulatory capacity of this transcription factor. This could happen, for example, if the binding of FHX to sequences of one type, but not to sequences of the other, expose signals that are potential targets of modulatory molecules.

The transactivation of the 5×FHX-A.Luc reporter system by FHX followed an exponential dose-response curve (inset in Fig. 7B). This is suggestive of a cooperative mode of activation of transcription by several FHX molecules. This cooperation could happen at the DNA-binding step. Since the FHX reporter construct used in these experiments contained five FHX sites, the occupation of one site could be favored if FHX molecules were already bound to nearby sites. This would result in a more than additive (synergetic) effect on transactivation. In contrast, the 5×Gal4.Luc reporter was activated by the chimeric factor Gal4DBD::FHX, which binds DNA independently of FHX sequences, following a linear dose-response curve (not shown), compatible with an additive effect on transcription. Cooperation at the transactivation step would still be possible; however, further experiments are required to address this issue.

Inspection of the FHX amino acid sequence did not show any clear potential transactivation domain. Nevertheless, a region within the C-terminal half of the molecule was reminiscent of glutamine-rich and proline-rich transactivation domains as those of Sp1 and CTF1/NF1 transcription factors, respectively (42, 43). This region has been conserved in the mouse homologue of FHX, which we have recently cloned, although the number of glutamine and proline residues was different (data not shown). This is probably because glutamines and prolines in this region are encoded by repeats of CAG and CCX triplets, respectively, the structure of which somehow resembles microsatellite DNA. If such, this region could undergo increments or decrements in the number of these triplets as described for other microsatellites as, for example, in the Huntington disease gene (44).

The ample tissue distribution of the FHX expression in the adult suggests that it could be important for some basic cellular function. This has been shown to be the case of Trident, also known as WIN or HFH-11, one of the fork head factors most related to FHX. Trident is involved in the correct synchronization of S phase and mitosis during the cell cycle (45). FHX function could also be important for the very early phases of the embryonic development. We found that, in the mouse development, the onset of FHX expression occurred very early, suggesting that the activation of the FHX expression occurred at, or immediately after, the zygotic gene activation, an important event for the first cell differentiation events of the embryonic development. It is tempting to speculate that the FHX function could be important in some zygotic genome-dependent phenomena that happen at these early stages of development.

**Acknowledgments**—We thank Luis Roiz and Josefa Fernández-Cabrera for excellent technical assistance and Dr. Mario García de Lacoza for the molecular modeling of the FHX fork head domain. We are grateful to Dr. Carlos López-Otín for critical readings and Dr. Santiago Lamas for wise advice for improving the manuscript. We are indebted to Dr. Augusto Silva for making available the facilities of his laboratory. We also thank Dr. J. Darnell for the generous gift of the HNF3α cDNA.

**REFERENCES**

FHX, a Novel Fork Head Factor with a Dual DNA Binding Specificity
Cristina Pérez-Sánchez, María Ana Gómez-Ferreriña, Carmen Arias de la Fuente,
Begoña Granadino, Gloria Velasco, Andrés Esteban-Gamboa and Javier Rey-Campos

doi: 10.1074/jbc.275.17.12909

Access the most updated version of this article at http://www.jbc.org/content/275/17/12909

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 46 references, 20 of which can be accessed free at
http://www.jbc.org/content/275/17/12909.full.html#ref-list-1