Constitutive Expression of Murine Complement Factor B Gene Is Regulated by the Interaction of Its Upstream Promoter with Hepatocyte Nuclear Factor 4*

(Received for publication, July 31, 1996)

Gerard Garnier, Antonella Circolo, and Harvey R. Colten†‡
From the Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri 63110

Factor B (Bf) is a constituent of the alternative pathway of complement activation encoded within the major histocompatibility complex. Transcription of the murine gene from two initiation sites generates two Bf mRNA species differing in size and tissue distribution. Striking genetic, tissue-specific differences in Bf mRNA levels at extrahepatic sites (kidney and intestine) among mouse strains correlate with a DNA sequence polymorphism in the 5′-flanking region of the gene and differential nuclear protein binding at the Bf upstream transcriptional initiation site (UIS). To ascertain the functional consequences of this polymorphism in the Bf promoter, we analyzed the effects of strain-specific sequences in the Bf 5′ region on the expression of a chloramphenicol acetyltransferase (CAT) reporter gene transfected in human and mouse hepatoma cells. The CAT activity and mRNA level produced when transcription was driven by the sequence of strains with high extrahepatic expression were reduced to background levels when the sequence specific to the low expres sor strains was used. Eighty percent of this difference was accounted for by a point substitution that affects DNA-protein interaction at the UIS, the sequence of higher affinity conferring higher expression. Hepatocyte nuclear factor 4 (HNF-4), derived from HepG2, mouse liver and kidney or cell-free translation of HNF-4 RNA, is the nuclear protein that preferentially binds to the high expressor UIS. Bf-CAT is not expressed in cells that lack HNF-4 (CV-1). However, co-transfection of HNF-4 into CV-1 cells drives Bf-CAT expression and reproduces the differences derived from the substitution that affect HNF-4 binding in vitro. These data show that interaction of HNF-4 with polymorphic variants of the upstream Bf promoter is the major determinant of strain-specific extrahepatic factor B expression.

Factor B (Bf) is a serine proteinase of the antibody-independent, alternative pathway of complement activation, an important humoral system of host defense against invading pathogens (1). In addition, Bf activation fragments exert cytokine-like activities such as B lymphocyte proliferation and differentiation (2–4), macrophage spreading (5), and monocyte-mediated cytotoxicity (6–8). The Bf gene is located within the class III gene cluster of the major histocompatibility complex (MHC) on chromosome 6 in humans (9) and 17 in mice (10, 11) immediately downstream of the complement C2 gene (12, 13), its structural and functional homologue in the classical, antibody dependent complement activation pathway.

The major site of Bf biosynthesis is the liver, as evidenced by allotypic changes of serum Bf following liver transplant (14). Hepatic constitutive Bf expression has been demonstrated in vitro, in primary hepatocyte cultures (15), in the HepG2 hepatoma cell line (16), and in vivo, by Bf mRNA analysis in liver tissue (17, 18). Bf is a positive acute phase reactant in that its hepatic synthesis and serum level are increased during the acute phase of the inflammatory response (19).

Extrahepatic biosynthesis of Bf and other complement components has also been documented in cell lines of nonhepatic origin and in vivo (19, 20). Extrahepatic Bf expression is usually low constitutively, but it is markedly up-regulated under inflammatory conditions by bacterial endotoxin and by cytokines such as interleukin-1, interleukin-6, tumor necrosis factor-α, and interferon-γ in many cell lines and tissues (17, 18, 21–25). The extrahepatic production of complement may be of biological importance especially at early stages of inflammation. In the mouse, the kidney is by far the major extrahepatic site of Bf mRNA constitutive expression (18, 25), although Bf mRNA is constitutively present to some extent in many other nonhepatic tissues.

Two mRNAs differing in size (2.4 and 2.7 kb) are transcribed from the murine Bf gene as a result of alternative transcriptional initiation at two sites separated by 302 bp (18). The downstream transcriptional initiation site (DIS) is located 105 bp 5′ to the Bf open reading frame at a position homologous to that of the human gene (12). The upstream initiation site (UIS) is only 88 bp 3′ to the polyadenylation site of the C2 gene (18). The two Bf mRNA species are present in equal amounts in kidney or intestine, but in the liver, the short mRNA represents at least 95% of Bf transcription products (18, 26). Bf up-regulation during inflammation is accounted for by a preferential increase of expression of the short transcript (18, 27, 28), which in turn is translated at approximately twice the rate of the long transcript (29). Thus, extrahepatic Bf expression is up-regulated by the cooperative effects of transcriptional and translational controls. Although the C2-Bf intergenic region is highly conserved between mouse and human (12), a longer Bf mRNA has not yet been found in humans.

In certain inbred strains (e.g. H-2× and H-2, MHC haplotypes), both Bf mRNA species are expressed at very low levels in the kidney and intestine, compared with other strains (28). This phenotypic difference correlates with DNA sequence polymorphism in the 5′ region of the Bf gene that includes the...
disruption of a binding site for liver and kidney nuclear protein(s) at the UIS in the low expressor strains (28). Hepatic expression of Bf in these strains is similar to that in other strains including those with high extrahepatic Bf expression.

Cis elements required for Bf up-regulation by cytokines have been mapped in the 5′-flanking regions of the mouse (12, 30) and human (12, 13) genes. However, the molecular mechanisms and transcription factors involved in the control of constitutive and tissue-specific expression of the Bf gene have not yet been identified. In this study, we demonstrate the functional relevance of the sequence polymorphism in the 5′ region of the murine Bf gene in Bf transcription. Differential interaction of the upstream promoter with HNF-4, resulting from the point substitution near the UIS, appears to be the major determinant in extrahepatic strain-specific Bf expression.

MATERIALS AND METHODS

Bf-CAT Reporter Constructs—Bf-CAT constructs were prepared and analyzed according to standard cloning and sequencing procedures (31) using Promega enzymes (Promega Corp., Madison, WI) and a Sequenase version 2.0 kit (U.S. Biochemical Corp.). Oligonucleotides were synthesized by a PCR-Mate DNA synthesizer (Applied Biosystems, Foster City, CA). Polymerase chain reactions were performed in a Temp-Cycler (Coy Laboratory Products, Ann Arbor, MI) using Perkin-Elmer AmpliDNA DNA polymerase and reagents and dNTP-Mixed (inc., Branchburg, NJ). Bf genomic clones (28) from B10.PL (high expressor) and B10.M (low expressor) mouse strains were used as DNA templates to generate strain-specific Bf 5′-flanking sequences by PCR. The plasmid pCAT-0, a subclone of pSV5-CAT in pBluescript-II SK(+) (32) (from which the EcoRI site in the 3′ polynucleotide was deleted) was used as PCR template for the CAT sequence and as vector for subcloning.

Strain-specific 5′ sequences of the murine Bf gene, extending from −728 to +105, the positions preceding the first in-frame ATG (numbering refers to the nucleotide positions relative to the DIS), were linked to the UIS and the 3′-UT end to the first 214 bp of the CAT open reading frame by overlap extension PCR (33). The chimeric PCR products, containing an engineered XbaI 5′ cloning site and the internal EcoRI site of the CAT gene at the 3′ end, were then ligated with the corresponding sites of pCAT.0, yielding the BF-CAT constructs pBFL-CAT (high expressor) and PBIM-CAT (low expressor). A mutant pBf-CAT.M1 was generated (34) by substituting T to C in the sequence of the low expressor strains when the sequence of the high expressor strains was used (pBfPL-CAT) did not significantly differ from background (pCAT.0).

The differences between high and low expressor sequences consists of six point substitutions and a 6-bp deletion. These sequences were analyzed in the upstream promoter and 5′-flanking regions by PCR. The plasmid pBfPL-CAT, a subclone of pSV5-CAT in pBluescript-II SK(+) (32) (from which the EcoRI site in the 3′ polynucleotide was deleted) was used as PCR template for the CAT sequence and as vector for subcloning.

Strain-specific 5′ sequences of the murine Bf gene, extending from −728 to +105, the positions preceding the first in-frame ATG (numbering refers to the nucleotide positions relative to the DIS), were linked to the UIS and the 3′-UT end to the first 214 bp of the CAT open reading frame by overlap extension PCR (33). The chimeric PCR products, containing an engineered XbaI 5′ cloning site and the internal EcoRI site of the CAT gene at the 3′ end, were then ligated with the corresponding sites of pCAT.0, yielding the BF-CAT constructs pBFL-CAT (high expressor) and PBIM-CAT (low expressor). A mutant pBf-CAT.M1 was generated (34) by substituting T to C in the sequence of the low expressor strains when the sequence of the high expressor strains was used (pBfPL-CAT) did not significantly differ from background (pCAT.0).

The differences between high and low expressor sequences consists of six point substitutions and a 6-bp deletion. A mutant (pBf-CAT.M1) was also generated, in which only the nucleotide at −299 (3 bp upstream of the UIS) of the high expressor sequence was substituted (T to C). This substitution disrupted the site for nuclear protein binding at the UIS (28). The CAT activity for this mutant was approximately 20% of the CAT activity obtained with the high expressor, indicating that 80% of the difference in CAT activity between the high and low expressor sequences results from this T/C substitution. Because the differences in Bf-CAT expression in HepG2 were more comparable with extrahepatic than hepatic differences in Bf expression in vitro, the following experiments were done. Analysis of expression of several constructs extending up to 6.7 kb upstream of the DIS (26) in HepG2 showed a correlation of CAT activity with sequences identical to that observed with the shortest constructs (see above). The longest (−6.7 kb) and shortest (−728 bp) constructs transfected into a murine hepatoma cell line (Hepa1–6), also yielded results similar to that observed in HepG2 transfected with the shortest Bf strain-specific sequences; i.e. expression was markedly reduced for the UIS substitution (72% and 78% reduction for the short and 5′-extended constructs, respectively).

Expression of Bf-CAT Reporter Constructs in Transfected Cells—Human (HepG2) and murine (Hepa1–6) hepatoma cells (American Type Culture Collection (ATCC), Rockville, MD), grown in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum, were plated at 2 × 10⁵ cells in 100-mm tissue culture dishes. The cells were transfected the next day with the calcium phosphate method (35) by a 7-h incubation with 10 μg of test plasmid and 3.5 μg of β-galactosidase expression vector pH110 (Pharmacia Biotech Inc.). The CAT activity was determined 36 h later by a modification (32) of a previously described method (36) in the cell lysates normalized for β-galactosidase activity. For transcriptional activation experiments using an HNF-4 expression plasmid (see below), monkey kidney-derived, fibroblast-like CV-1 cells (ATCC) were plated at 10⁵ cells/dish and co-transfected with 6 μg of CAT plasmid, various amounts (see “Results”) of HNF-4 plasmid, and 3.5 μg of β-gal plasmid, processed, and analyzed as above.

For Northern blot analysis of CAT and β-galactosidase mRNAs, total cellular RNA was prepared by guanidinium thiocyanate lysis and CsCl density gradient ultracentrifugation (37), and 20 μg were separated in a 1.2% agarose denaturing gel and analyzed as described (28). The CAT cDNA probe was a 516-bp NcoI fragment corresponding to the most 5′ CAT coding sequence from the pBfM-CAT plasmid described above. The β-galactosidase cDNA probe was a 2.4-kb PvuII fragment from pH110. RNA loading and quality were controlled by ethidium bromide stain of ribosomal RNA in the gel and on the blot.

Electrophoretic Mobility Shift Analysis—Nuclei from mouse tissues and HepG2 cells were prepared as described (38) and lysed with NaCl and spermidine (39), and the nuclear lysates were dialyzed (40). The soluble material was quantified for protein content using a protein assay kit at 70°C. The binding reactions were performed essentially as described (29) in the presence of 4 μg of extract, 2 μg of poly(I·dC)-poly(dI·dC), and 175 pg of 32P-end-labeled, 35-mer double-strand oligonucleotide as probe in 25 μl containing 5% glycerol, 1 mM EDTA, 1 mM di-thiothreitol, 1 mM phenylmethylsulfonyl fluoride, 50 mM NaCl, 10 mM Tris-HCl, pH 7.5, for 30 min at room temperature. The complexes were resolved in a 7% polyacrylamide gel in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0).

The sequences of the probes are indicated under “Results” except for the CHIB site of the rat apoCIII gene (−1039 to −69, coding strand: 5′-TCTGGAGCTTGTCGACGTTCCCTTGGACACGTC-3′) (41) and the 5′-UT binding site of the mouse Bf gene (+368 to +70, coding strand: 5′-CAGGAGGTCAATAGGCCCCAGATGACGTGAG-3′) (12). The latter sequence is compatible with a 728- to 116-bp genomic probe in which the binding element of the 5′-UT region was initially detected (28) and therefore contains the complete 5′-UT element (not shown).

RESULTS

A Polymorphic Binding Site in the Bf Upstream Promoter Responsible for Strain-specific Constitutive Activity of the Downstream Promoter—To ascertain the functional consequences of the sequence differences in the 5′ region of the mouse Bf gene between strains with high and low extrahepatic Bf expression, we generated a set of CAT reporter constructs driven by the strain-specific Bf 5′ sequences (−728 to +108 from the DIS) (Fig. 1B), transfected them into HepG2 cells, and assessed the CAT activity (Fig. 1A). Higher activity was obtained with the construct bearing the high expressor strain sequence (pBFL-CAT). In contrast, CAT activity detected when the sequence of the low expressor strains was used (pBIM-CAT) did not significantly differ from background (pCAT.0). The differences between high and low expressor sequences consists of six point substitutions and a 6-bp deletion. A mutant (pBf-CAT.M1) was also generated, in which only the nucleotide at −299 (3 bp downstream of the UIS) of the high expressor sequence was substituted (T to C). This substitution disrupted the site for nuclear protein binding at the UIS (28). The CAT activity for this mutant was approximately 20% of the activity obtained with the high expressor, indicating that 80% of the difference in CAT activity between the high and low expressor sequences results from this T/C substitution. Because the differences in Bf-CAT expression in HepG2 were more comparable with extrahepatic than hepatic differences in Bf expression in vivo, the following experiments were done. Analysis of expression of several constructs extending up to 6.7 kb upstream of the DIS (26) in HepG2 showed a correlation of CAT activity with sequences identical to that observed with the shortest constructs (see above). The longest (−6.7 kb) and shortest (−728 bp) constructs transfected into a murine hepatoma cell line (Hepa1–6), also yielded results similar to that observed in HepG2 transfected with the shortest Bf strain-specific sequences; i.e. expression was markedly reduced for the UIS substitution (72% and 78% reduction for the short and 5′-extended constructs, respectively).

Differences in CAT gene expression were confirmed at the mRNA level by Northern blot analysis of the CAT mRNA in transfected HepG2 cells (Fig. 1C). A CAT transcript was also
detected with an upstream construct (pUBiPL-CAT) in which the CAT open reading frame was inserted at the first ATG following the UIS (22bp downstream of the UIS) (Fig. 1C, lane 6), indicating that the sequence spanning −272 to −276 contains functional element(s) of the Bf upstream promoter. However, no significant CAT activity was detected using this construct (not shown), probably due to low translation initiation at this ATG (29).

To analyze the effect of strain-specific sequences on the activity of the upstream promoter, constructs bearing the upstream sequence of the low expressor strains (pUbM-CAT) and identical sequence substituted only at −299 (pUbF-CAT.M1) were also generated (Fig. 2A) and transfected into HepG2 cells, and the expression of the CAT transcript was assessed by Northern blot analysis (Fig. 2B). Transcription driven upstream by the low expressor sequence was also markedly less than that driven by the high expressor sequence. However, in contrast to downstream initiation, the substitution at −299 (T to C) had no effect on transcription initiated at the UIS. A 5′-truncated construct (pUPbF-CAT) retained the promoter activity, which thus localizes the minimal upstream promoter within a 272-bp fragment (−553 to −276 from the DIS).

The differential binding at the UIS resulting from the T/C polymorphism at −299 (28) was further analyzed by gel shift analysis using 35-mer double-strand oligonucleotides (−323/−289) as probes (Fig. 3). Using a mouse liver nuclear extract, the high expressor sequence as probe (T at −299) and increasing amounts of competitor with or without the substitution (T to C) at −299, the gel shift analysis (Fig. 3A) shows that both sequences bind to the nuclear factor(s), but approximately 8 times more competitor was required to reach equivalent competition when T was substituted to C. Single-strand substitutions in the coding strand did not affect the interaction with the mouse liver and kidney and HepG2 nuclear protein(s). In contrast, the corresponding A to G substitution in the noncoding strand abrogated the interaction with all three ex-
tracts. The HepG2 extract showed an additional complex of faster mobility. However, this interaction was consistently decreased only by the single-strand substitution in the coding strand and equally competed by the wild-type and double-strand mutant oligonucleotides (not shown).

The Two Major Binding Sites at the UIS and in the 5'-UT of the Bf Gene Share Transcription Factors with the Apolipoprotein CIII Gene Promoter—The binding specificity at the UIS and in the 5'-untranslated region of the Bf gene was assessed by competition with a sequence of the apolipoprotein CIII gene that shows homologies with the Bf sites and is known to interact with members of the steroid hormone receptor superfamily (42, 44, 45). Double-strand oligonucleotides were generated for the UIS (−323/−289) and 5'-UT (+36'/+70) sites of Bf and for the CIIB site (−103'/−69) of the rat apoCIII gene. These were used as probes and competitors in gel shift analysis with mouse liver and HepG2 nuclear extracts (Fig. 4). All three probes showed specific interactions with both extracts (Fig. 4A). With the mouse liver extract, the two Bf sites did not compete with each other, but each was competed by the CIIB site. The CIIB site was only partially competed by each of the two Bf sites separately but completely competed when both Bf sequences were added (Fig. 4B). The kidney nuclear extract showed identical results (not shown). With the HepG2 extract an additive effect of the Bf competitors on the CIIB interactions was also observed (Fig. 4A). As with the liver and kidney extracts, the interaction with the UIS probe that was affected by the T/C polymorphism at −299 was competed by the CIIB site but not by the 5'-UT site (Fig. 4A).

The T/C Polymorphism at −299 Affects the Affinity of the UIS Element for HNF-4—Sequence homologies and binding specificities of the UIS and elements interacting with members of the steroid hormone receptor superfamily and the tissue distribution of these factors led us to analyze the interaction of HNF-4 with the binding site at the UIS. An HNF-4 transcript synthesized in vitro from a rat cDNA clone (42) was translated in a reticulocyte lysate. The interaction of the translation product with the UIS element was assessed by gel shift analysis using as probe the high expressor sequence (−323/−289) described above. The kinetics of translation (Fig. 5A) and the dose

**Fig. 3.** The T/C substitution near the UIS affects the affinity of a DNA element for a nuclear factor. A, differential binding of the two alleles. Gel shift analysis was performed using the wild-type (T at −299) double-strand oligonucleotide as probe and a mouse liver nuclear extract in the presence of increasing amounts (25-, 50-, 100-, and 200-fold excess) of wild type or mutant (C at −299) competitor. B, the substitution on the noncoding strand affects the interaction at the UIS. Wild-type, mutant, and hybrid (with single-strand substitutions) double-strand oligonucleotides with the residue at −299 as indicated were used as probes in gel shift analysis with the indicated nuclear extracts. All interactions were competed by the wild-type competitor (not shown). C, oligonucleotides (coding strand is shown only) used in A and B, showing the T/C polymorphic residue at −299 (highlighted).

**Fig. 4.** Competition for binding between the DNA elements at the UIS and in the 5'-UT region of the Bf gene and the CIIB site of the apoCIII promoter. A, gel shift analysis of the interactions of the Bf UIS and 5'-UT elements and the CIIB site of the apoCIII gene with mouse liver and HepG2 nuclear proteins. Competitions were performed with a 50- and 100-fold excess of the indicated competitor. B, competition of the interactions at the CIIB site by the Bf UIS and 5'-UT binding elements added separately (50-, 100-, and 200-fold excess) and together (50- and 100-fold excess of each) to the reactions with mouse liver (upper part) and HepG2 (lower part) nuclear extracts. The six control lanes on the left show the interactions on each probe and competition with a 100-fold excess of the corresponding cold probe (CP).
response to HNF-4-specific RNA input (Fig. 5B) show formation of a complex between the UIS probe and cell-free translated HNF-4. The complex with HNF-4 had the same mobility as the complex obtained with the mouse nuclear extract, was not competed by the 5'-UT site, and was competed by the CHIP site of the apoCIII gene (Fig. 5C). The substitution (T to C at −299) specific for the low expressor sequence markedly reduced the HNF-4 interaction at the UIS as shown by using the mutated probe and competitor (Fig. 5C, lanes 8 and 5, respectively). A complex of slightly higher mobility with a protein already present in the lysate was also detected.

The complex formed on the UIS probe with the mouse liver nuclear extract was recognized by an HNF-4-specific antiserum (42), as assessed by supershift analysis (Fig. 6A). The specificity of the supershift complex is indicated by the lack of such supershift using a nonimmune serum, by the absence of direct interaction between the antisera and the probe, and by the identical supershift pattern obtained with cell-free translated HNF-4. Similarly, the complexes formed with the kidney and HepG2 nuclear extracts were recognized by the antiserum (Fig. 6B). The faster mobility complex with the HepG2 extract was not shifted by the antibody.

**Strain-specific Transcriptional Activation of Bf-CAT Constructs by Exogenous HNF-4 in Transfected CV-1 Cells**—To confirm a functional role for HNF-4 in strain-specific Bf transcription, we cotransfected CV-1 cells, a monkey kidney-derived cell line that lacks endogenous HNF-4 (46), with the aforementioned Bf-CAT constructs (Bf 5' sequences: −728 to −108 from the DIS and an HNF-4 expression plasmid. Cotransfection with HNF-4 increased Bf-CAT expression in a dose-dependent manner (Fig. 7). At 0.1 μg of HNF-4 plasmid input (and up to 0.6 μg, not shown), transcription from the high expressor Bf sequence was markedly greater than with the low expressor. Most of this effect was accounted for by the mutation at the UIS (T to C at −299) in a manner similar to that observed in hepatoma cells HepG2 (Fig. 1), which express endogenous HNF-4. However, at higher HNF-4 plasmid input, (4 μg), the T to C substitution had little effect on Bf-CAT expression. On the other hand, expression driven by the low expressor sequence was significantly less, even at high HNF-4 input.

**A Low Affinity Binding Site for HNF-4 in the Human Bf Gene**—The HNF-4 antibody supershift analyses described above show that in the mouse genomic sequence, the cis element encompassing the UIS interacts with both murine and human HNF-4. The high degree of sequence homology between the 5'-flanking regions of the human and murine Bf genes (12), particularly in the sequence surrounding the UIS, raises the question whether HNF-4 also interacts with the corresponding human sequence. In the human sequence, the residue corresponding to the position of the T/C polymorphism that affects HNF-4 binding in mice (−299 in the mouse, −297 in the human) is a cytosine (Fig. 8C) as in the low expressor strains. To compare the DNA-protein interactions between the murine and human sequences, we used the mouse UIS probe described above (−323 to −289, high expressor sequence) and its human counterpart (−321 to −287) in gel shift analysis. With the mouse liver and kidney and HepG2 nuclear extracts, the two probes formed complexes of same mobility and competed with each other, although the interaction with the human sequence was markedly weaker in all cases (Fig. 8A). As for the mouse, the interaction of the human probe with the mouse liver and HepG2 extracts was dramatically increased by a C to T substitution at −297 (Fig. 8B, lanes 1 and 2, and lanes 4 and 5). The resulting complex was competed by the mouse sequence (high expressor) (lane 3) and recognized by the HNF-4 antiserum.
Transcriptional Control of Factor B Gene by HNF-4

We previously reported that differences in extrahepatic levels of factor B mRNA among mouse strains correlate with a DNA sequence polymorphism in the 5' region of the Bf gene and with binding of nuclear protein(s) expressed in liver and kidney (28). In the present study, we show that this polymorphism similarly affects the expression of a reporter gene in vitro (i.e. is of functional importance) and that the functional difference is primarily accounted for by a point substitution near the Bf upstream transcriptional initiation site that also markedly affects the affinity of a binding site for HNF-4 and transcriptional activation by HNF-4. The remaining differences between transcriptional activities of high and low expressor sequences (including a difference in the upstream promoter activity) originate from other polymorphic differences within this region.

In vivo Bf expression varies among various mouse strains particularly in extrahepatic sites, but only minimal differences in hepatic expression have been observed (17, 18, 28). The functional analysis of the murine Bf promoter was initially undertaken by transfecting Bf reporter (CAT) constructs into the well differentiated human hepatoma HepG2. Surprisingly, expression of Bf-CAT in HepG2 reflected more closely the sequence requirements for transcriptional initiator activity (47, 48) (except for A at +1 and T at +3). Substitutions in the region of the upstream transcription start site of the low expressor strains markedly reduce the upstream promoter activity in vitro, consistent with the low extrahepatic expression of the long Bf mRNA observed in these strains in vivo (28). One of these substitutions is located between two imperfect half-palindromes (CAGGAGG and CCTTGCTG) at −25 from the UIS. The sequence surrounding −30 has been shown to be important for transcriptional initiation, even in TATA-less promoters (49). The substitution of T to C in the HNF-4 site alone does not significantly affect the upstream promoter activity in transfected HepG2 cells. However, the interaction of HNF-4 with a transcriptional start site is unusual and suggests a possible role in the enhancement of the upstream promoter activity in conditions, like those prevailing in the kidney, that differ from HepG2. Although it has not been documented for HNF-4, several members of the steroid hormone receptor superfamily, including or-

DISCUSSION

Although it has not been documented for HNF-4, several members of the steroid hormone receptor superfamily, including or-

Fig. 7. Effects of strain-specific Bf 5'-flanking sequences on transcriptional activation of Bf-CAT expression by exogenous HNF-4. CV-1 cells were cotransfected with the indicated Bf-CAT construct (Fig. 1), 0.1 (□) or 4 µg (■) of the HNF-4 expression plasmid, or an equivalent amount of vector (□) as control and the β-galactosidase expression plasmid (see “Materials and Methods”). The relative CAT activity is normalized to the activity of the pBfPL-CAT construct without HNF-4 and represents the average values of two to five separate experiments.

(lane 6). Furthermore, the mutated human probe formed a complex with cell-free translated HNF-4 (lane 8).

Fig. 8. Comparison between the UIS and the corresponding sequence of the human gene for HNF-4 binding. A, interactions with mouse liver and kidney and HepG2 nuclear proteins and competition between the human (h) and mouse (m) sequences. B, C to T substitution at −287 in the human sequence increases HNF-4 binding. Gel shift analysis with the wild-type (lanes 1 and 4) and mutated (lanes 2, 3, and 5–8) human probes was performed with mouse liver (lanes 1–3) and HepG2 (lanes 4–7) nuclear extracts and cell-free translated HNF-4 (lane 8) alone or in the presence of mouse competitor (lane 3), HNF-4 antisera (lane 6), or nonimmune serum (lane 7). C, oligonucleotides used in A and B and comparison between the murine (high expressor) and human sequences; the polymorphic residue at −299 in mice and the corresponding residue in humans are highlighted.
Transcriptional Control of Factor B Gene by HNF-4

The sequence surrounding the USI consists of two 6-bp palindromic repeats (GATGGA and TCCATC, which are parts of a larger imperfect palindrome) separated by a 6-bp spacer containing a typical CTTTTG motif found in HNF-4 binding sites of other genes (45, 52, 53). Palindromic sequences are also found in some of these sites (52, 53). The consensus sequence for HNF-4 binding is not clearly established yet, due to the high sequence variability among the known HNF-4 binding sites, variability in their affinity for HNF-4, and in most cases their ability to interact with other proteins. Mutations within the CTTTTG motif invariably disrupt the site and decrease transcription (52, 54, 55, and this work). In the human blood coagulation factor IX gene, a substitution of the second T of the spacer (as in the mouse Bf gene) leads to developmentally determined factor IX deficiency and hemophilia B Leyden (55). In the Bf upstream promoter, at this position the double-strand substitution or single-strand substitution in the noncoding strand, but not in the coding strand alone, affects HNF-4 binding. The fact that a mismatch even at a critical position for binding does not impair the interaction suggests that in the complex with HNF-4 the two DNA strands may be locally dissociated. A strand separation at the upstream transcription start site, stabilized by intrachain associations between the half-palindromes and perhaps by the HNF-4 interaction, could facilitate promoter melting and hence transcription.

The high sequence homology between the human and mouse Bf 5'-flanking regions (12) led us to identify an HNF-4 binding site in the human gene at a position equivalent to that of the mouse sequence. A genomic DNA segment containing this site has been shown to function as a typical constitutive, cell-specific enhancer for human Bf transcription (13). This region, which is also important for transcriptional termination of the C2 gene, includes a binding site for the myc-associated zinc finger protein (56) and a 100-bp sequence further downstream (encompassing the HNF-4 site); both are required for accurate C2 transcriptional termination (56, 57). Identification of DNA/protein interactions in this region is therefore important for further analyses of the C2-Bf intergenic region. The interaction between HNF-4 and the human sequence is, however, relatively weak, due to a cytosine at −297, which in the mouse sequence characterizes the binding site of the low expression strains. However, as discussed above, a weak interaction may not preclude but rather restrict HNF-4-regulated expression of the human Bf gene to tissues such as the liver, where HNF-4 is abundant. This should prompt a search for polymorphism in the human sequence as well.