Factor VII is a vitamin K-dependent coagulation protein essential for proper hemostasis. The human Factor VII gene spans 13 kilobase pairs and is located on chromosome 13 just 2.8 kilobase pairs 5′ to the Factor X gene. In this report, we show that Factor VII transcripts are restricted to the liver and that steady state levels of mRNA are much lower than those of Factor X. The major transcription start site is mapped at −51 by RNase protection assay and primer extension experiments. The first 185 base pairs 5′ of the translation start site are sufficient to confer maximal promoter activity in HepG2 cells. Protein binding sites are identified at nucleotides −51 to −32, −63 to −58, −108 to −84, and −233 to −215 by DNase I footprint analysis and gel mobility shift assays. A liver-enriched transcription factor, hepatocyte nuclear factor-4 (HNF-4), and a ubiquitous transcription factor, Sp1, are shown to bind within the first 108 base pairs of the promoter region at nucleotide sequences ACTTTG and CCCCCCC, respectively. The importance of these binding sites in promoter activity is demonstrated through independent functional mutagenesis experiments, which show dramatically reduced promoter activity. Transactivation studies with an HNF-4 expression plasmid in HeLa cells also demonstrate the importance of HNF-4 in promoting transcription in non-hepatocyte derived cells. Additionally, the sequence of a naturally occurring allele containing a previously described decanucleotide insert polymorphism at −323 is shown to reduce promoter activity by 33% compared with the more common allelic sequence.

*This work was supported by National Institute of Health Grant HL48322 (to K. A. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EMBL Data Bank with accession number(s) U40852.

† To whom correspondence should be addressed: Abramson Research Center, Rm. 310A, 34th St., and Civic Center Blvd., Philadelphia, PA 19104. Tel.: 215-590-4522; Fax: 215-590-4834.

‡ The abbreviations used are: HNF-4, hepatocyte nuclear factor-4; bp, base pair(s); kb, kilobase pair(s); PCR, polymerase chain reaction; PIPES, 1,4-piperazinediethanesulfonic acid.
EXPERIMENTAL PROCEDURES

Genomic Library Screening, Subcloning, and Sequencing—A PCR fragment generated from the 279 bp immediately upstream of the Factor X translation start site was used as a probe to screen a human cosmid library as described previously (7). This cosmid clone contains 3.5 kb of the Factor X gene, the 2.8-kb Factor VII–Factor X intergenic region (8), the entire 12.8-kb Factor VII gene and approximately 10–15 kb of the Factor VII 5′-flanking sequence.

DNA Sequencing—The 5′-flanking DNA generated by EcoRI cleavage between base pairs –309 and –3958 of the human Factor VII gene was subcloned into a pUC19 vector for DNA sequence analysis. Oligonucleotide sequencing primers (Integrated DNA Technologies) were synthesized at approximately 300-bp intervals, and sequencing was performed using 7-deaza-dGTP termination mixtures (U.S. Biochemical Corp.) to minimize compression bands using the dye-exchange Sanger sequencing method (25). All sequences were confirmed on both strands. The sequence was scanned for putative transcription factor binding sites using version 2.3 of the TRANSFAC data base.2

Northern Blot Analysis—A 1.3-kb EcoRI-EcoRI fragment of Factor X cDNA and a 1.5-kb HindIII-EcoRI fragment of Factor VII cDNA were randomly labeled with [α-32P]dCTP using a Random Primed DNA labeling kit (Boehringer Mannheim) according to the specifications of the manufacturer, yielding probes with specific activities of 1.2 × 109 cpm/μg and 8 × 108 cpm/μg, respectively. On consecutive days, the Factor X and Factor VII cDNA probes were hybridized for 1 h at 65 °C in ExpressHyb hybridization solution (Clontech) with multiple tissue Northern blots containing approximations of human poly(A)+ RNA from 16 different tissues (Clontech). Densitometric measurements were performed on autoradiograms exposed to the membrane for 12 h using a Molecular Dynamics densitometer and ImageQuant™ software.

Primer Extension Assay—Primer extension was performed using an end-labeled 22-bp oligonucleotide (+21 to +42). Total human liver RNA was isolated from discarded segments of a donor transplant specimen using an RNA extraction kit (Pharmacia Biotech Inc.). One million cpm of primer was added to 50 μg of total human liver RNA and incubated for 90 min at 65 °C. The RNA was reverse transcribed for 1 h at 42 °C using 8 units of avian myeloblastosis virus reverse transcriptase (Promega) in a buffer containing 20 mM Tris-HCl (pH 8.3), 10 mM MgCl2, 0.5 mM dithiothreitol, 0.15 mg/ml actinomycin D, 14 mM 4-dNTP mix, and 50 mM NaCl. After reverse transcription, RNase A (2.1 μg) and 10 μg of salmon sperm DNA were added to the reaction in a total volume of 150 μl. The product was precipitated, phenol/chloroform extracted, resuspended in an 80% formamide loading solution, and analyzed on a 6% polyacrylamide gel alongside a sequencing reaction using [α-32P]dATP and the +21 to +42 bp probe as a sequencing primer. The primer extension was also performed with Superscript II (Gibco BRL/Life Technologies, Inc.) and poly(A)+ RNA using the +21 to +42 bp primer. The procedure was as described previously with the following minor changes. 1 μg of poly(A)+ human liver RNA was used in place of the total liver RNA, and the reverse transcription buffer contained the following: 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl2, 10 mM dithiothreitol, 4-dNTP mix, and 10 μg of poly(A)+ RNA was prepared using a PolyATtract® system 1000 kit (Promega). RNase Protection Assay—The transcriptional start site of the human Factor VII gene was analyzed using an RNase protection assay essentially as described previously (27). A 550-bp PCR-generated fragment spanning from –508 to +42 was subcloned into the SmaI site of pBluescript II SK (+) (Stratagene), and the correct sequence was verified to ensure that no polycloning site-induced errors were introduced. The plasmid was linearized and transcribed with [α-32P]CTP using 10 units of T7 RNA Polymerase (Promega). Approximately 500 ng of the Factor VII riboprobe, with a specific activity of 1 × 108 cpm/μg, was dissolved in hybridization buffer (80% formamide, 40 mM PIPES (pH 6.4), 0.4 mM NaCl, 1 mM EDTA), denatured with 10 μg of total liver RNA and denatured for 5 min at 85 °C before hybridization for 16 h at 45 °C. The hybridized samples were then incubated for 30 min at 30 °C with either RNase T1 alone (2 μg/ml) or RNase T1 (2 μg/ml) + RNase A (10 μg/ml). The positive control, pTRI-β-actin (Ambion), a 249-bp β-actin fragment cloned into the pTRIPLEScript vector, was hybridized with 10 μg of human liver RNA under the same conditions described above for the Factor VII probes. The negative control was the Factor VII probe hybridized with 10 μg of RNA (Life Technologies, Inc.) under identical conditions. The RNA Century Marker template set (Ambion) was used to generate 100-bp RNA markers using the protocol described by the manufacturer. The distance migrated by the Factor VII product on a 6% polyacrylamide gel was used to calculate the nucleotide length of the protected fragment based on an equation generated by plotting the known RNA marker size versus distance traveled using semilog analysis (28).

Construction of Growth Hormone Reporter Gene Plasmids—The promoterless plasmid pOGH (Nichols Diagnostics Institute), a 4.187-kb HindIII–EcoRI fragment of 5′ flanking sequences of the human growth hormone gene, was used for reporter gene plasmid constructs (29). All inserts were cloned into the multiple cloning site between the BamHI and HindIII sites, spaced 17 bp apart. The 3′-end of all inserts ended at position +1, the first nucleotide of the translation start site (6). The 75-, 186-, 291-, 404-, and 501-base pair inserts were synthesized using PCR products with HindIII and BamHI sites artificially introduced at 5′ and 3′-ends, respectively. The PCR fragments were subcloned into the pOGH vector by directional cloning. The 109-bp constructs (wild-type and mutant) and the 162-bp construct were generated using a 3′-end primer with an artificially installed BamHI site and 5′-end primer containing sites sites using version 2.3 of the TRANSFAC data base.2

2 Wingender, E. (1994) www. transfac.gbf-braunschweig.de/ index.html.
of pOGH containing 309 bp of Factor VII (−309 to −1) was digested with BamHI, dephosphorylated by calf intestine alkaline phosphatase (Life Technologies, Inc.), end-labeled with [γ32P]ATP, and subsequently digested with HindIII. Approximately 10 ng of probe (6 × 10^4 cpm) was incubated for 30 min on ice with 50 μg of nuclear extract in 40 μl of buffer (25 mM Hepes (pH 7.6), 60 mM KCl, 7.5% glycerol, 0.1 mM EDTA, 0.75 mM dithiothreitol, 5 mM MgCl2, and 1 μg of poly(d-dC) (Pharmacia), and 1 μg of salmon sperm DNA, in 20 μl of binding buffer (25 mM Hepes, pH 7.6), 14 mM NaCl, 10% glycerol, 0.1 mM EDTA, 0.75 mM dithiothreitol, 5 mM MgCl2) before 10^7 cpm of DNA probe was added. Samples were incubated on ice for 30 min. In competition studies, unlabelled oligonucleotide was added during the preincubation step. The Sp1 consensus oligonucleotide (Promega) contains the following sequence: 5′-ATTCCGATGCGGGGCGGAGC-3′. In supershift experiments, antibody was added after the 10-min incubation period of the probe with the nuclear extract mixture. The Sp1 antibody (Santa Cruz Biotechnology) does not cross-react with Sp2, Sp3, or Sp4. The HNF-4 antiserum (rabbit) was raised against a peptide corresponding to amino acids 445–455 of the rat HNF-4 (gift from Dr. James Darnell and Dr. Frances Sladek, Rockefeller University). In studies with purified Sp1, one footprint unit of Sp1 (Promega) was used in place of the nuclear extract. Gel mobility shift assays were performed on 5% polyacrylamide gels (acylamide/bisacrylamide, 29:1) in 0.5 × Tris borate/EDTA for 2 h at 4 °C after prerunning the gel for 2 h at 200 V.

**RESULTS**

Northern Analysis of Factor VII Reveals Liver-specific Expression—Northern analysis of human poly(A)+ RNA revealed a single 2.7 kb band present only in liver (Fig. 1, arrow at left) using blots containing 16 different human tissue sources. Steady state Factor VII mRNA levels were 6% of those of Factor X mRNA levels, as determined by densitometric measurement of autoradiograms of the blot incubated sequentially with Factor VII and Factor X cDNA probes of comparable specific activity. These levels of steady state mRNA correlate well with molar plasma levels (12) of Factor VII (0.009 μM) and Factor X (0.18 μM).

DNA Sequence of the Human Factor VII Gene 5′-Flanking Region—The 5′-flanking sequence of the proximal promoter of the human Factor VII gene is shown in Fig. 2. Minor differences were observed between the first 521-bp sequence reported by O’Hara et al. (6) (isolated from HepG2 cells) and this sequence. In our sequence an extra C was seen at −149, and an extra G was seen at −459.

The TRANSFAC data base was used to search for putative transcription factor binding sites. Potential binding sites were

---

**FIG. 1. Northern analysis of mRNA from multiple human organs demonstrates liver-specific expression of human Factor VII mRNA.** A nylon membrane containing approximately 2 μg lane of poly(A)+ RNA was hybridized to a radiolabeled probe containing a 1.5-kb HindII-EcoRI cDNA fragment of the human Factor VII gene. The 2.7-kb band as indicated by an arrow corresponds to the Factor VII transcript. A β-actin control on the same blot permits normalization.

**FIG. 2. Nucleotide sequence of the proximal promoter of the human Factor VII gene.** The full 3958-base pair 5′-flanking DNA sequence can be found in GenBank (accession number U40852). The proximal 455 bp of sequence are derived from an allele, which does not contain the previously described decanucleotide insert at −323. The sequence from −3958 to −456 was obtained from an allele that contains the decanucleotide insert. The major transcription start site is marked with an asterisk at −51. The 5′-flanking region is shown in lower case. The first 42 bp in exon 1 are shown in capital letters with the corresponding single-letter amino acid code underneath. Protein-binding sites identified by DNase I footprint analysis and gel mobility shift assays are shaded. Putative binding sites identified within the first 400 bp are as follows: AP-1, GATA factors, Sp1, AP-4, TFEB, XPF-1 (exocrine pancreas factor-1), HRE (hormone response element), Sp1, GAGA factor, H4TF-1 (Histone H4 transcription factor), GR (glucocorticoid receptor), and HNF-4.
seen for the liver-enriched transcription factors HNF-4 at -67 to -56 and LF-A1 (Liver Factor-A1) (43) at -2178 to -2162. Putative binding sites were also seen for the following transcription factors at the indicated positions (using the numbering system of O'Hara et al. (6) where -1 represents the first nucleotide 5' to the translation start site): Sp1 at -286 to -276, -204 to -195, and -101 to -92; XPF-1 (exocrine pancreas factor-1) at -221 to -207; AP-1 at -391 to -382; AP-2 at -1684 to -1675; AP-4 at -3487 to -3478 and -234 to -225; GR (glucocorticoid receptor) at -91 to -82; H4TF-1 (histone H4 transcription factor-1) at -100 to -92; PTF-1 (pancreatic acinar cell-specific factor) at -1403 to -1392; GATA transcription factors at -363 to -353, TFE8 at -224 to -213; a hormone response element at -226 to -212; GAGA factor at -143 to -130, and ISGF-1 (interferon-stimulated gene factor-1) at -3552 to -3543. An Alu repeat is present at -956 to -634. Additionally, two nucleotide differences were noted between an individual containing a decanucleotide sequence insert (5'-CCTATATCCT-3') at -323 and the sequence of an individual lacking the insert; these changes were at -122 (T to C) and -401 (G to T). The sequence data beyond -456 is based on the sequence present in the cosmids clone, which was constructed from an individual in whom the previously described decanucleotide insert was present.

**Transcription Start Sites of Human Factor VII**—The start sites of transcription of the TATA-less human FVII gene were determined by primer extension and by RNase protection assays. Primer extension was performed using avian myeloblastosis virus reverse transcriptase and a primer located at +21 to +42 in exon 1 to reverse transcribe total human liver RNA (Fig. 3A). A cluster of start sites was seen after a 2-week exposure of the autoradiogram; the most prominent bands mapped to -71, -69, and -68 (upper triple arrow). Bands at positions -51, -50, and -48 (lower triple arrow) and -32 (single arrow) were also visible. Less distinct bands were seen around -105, -117, -130 (within the bracket at right) with the reverse transcribed product generated by avian myeloblastosis virus reverse transcriptase. The results were confirmed using liver poly(A)+ RNA and Superscript II primed with the +21 to +42-bp oligonucleotide (data not shown). In addition to bands seen at -71, -69, -68, -51, and -50, this second primer extension also showed bands at -38, -45, -79, and -80.

RNase protection was performed using total human liver RNA hybridized to a riboprobe (Fig. 3B). Results were obtained using digestion with RNase T1, which cleaves 3' to guanine residues, and with RNase T1 + RNase A, which together cleave
performed in triplicate. Mutation of the ACTTTG NF-4 binding motif (186 bp of the 186-bp wild-type construct). The data represent the mean ± S.D. of five independent experiments in HepG2 cells and three to six transfections from two independent experiments in HeLa cells. For purposes of comparing promoter activity in HepG2 and HeLa cells, the pOGH negative control in HeLa cells was normalized to the HepG2 negative control. The standard deviations in HeLa cell experiments ranged between 3 and 30% for all constructs; error bars cannot be seen with HeLa transfections in this graph due to the low values. B, reporter gene constructs (109-bp constructs) containing mutations at −100 and −94 in the G-rich region of the Sp1 motif (CCCCCTCCCCC mutated to CCACCTCCAC) were made as described under “Experimental Procedures.” The normalized promoter activity of a wild-type construct and three mutant constructs (109 bp MT94/100: C to A changes at −94 and −100; 109 bp MT94: C to A change at −94 only; and 109 bp MT100: C to A change at −100 only) are shown as a percentage of the 186-bp wild-type construct. The data represent the mean ± S.D. based on two independent transfection experiments in HepG2 cells performed in triplicate. Mutation of the ACTTTG HNF-4 binding motif (186 bp, ACTTTG) between bases −63 and −58 to GACAAT was generated within the 186-bp construct. The data represent the mean ± S.D. based on four transient transfections of HepG2 cells.

3′ to guanine and pyrimidine residues. The product incubated with RNase T1 alone revealed a prominent band of 95 bp, which maps to a transcription start point at −53 (lane 4, upper arrow on right). This was two to three base pairs longer than the most prominent RNase T1 + RNase A digested products, which yielded two adjacent bands corresponding approximately to positions −51 and −50 (lane 5, lower arrows on right). The positive control, a 249-bp β-actin riboprobe, was seen at the appropriate size (arrow on left) with the RNase T1-digested product (lane 2) but was again several base pairs longer than the RNase T1 + RNase A product (lane 1). No bands were seen in the negative control lane (lane 6) using tRNA as a source for hybridization with the Factor VII probe.

Taken together, the primer extension and RNase protection data suggest a major start site at approximately −51. The primer extension data also suggest the presence of start sites between −80 and −30, although these were not confirmed by RNase protection.

Factor VII Promoter Activity in HepG2 and HeLa Cells—To determine the fragments of the Factor VII 5′-flanking region required for maximal promoter activity, segments were cloned into a reporter gene plasmid containing the human growth hormone structural gene and transfected into either HepG2 cells or HeLa cells. HepG2 cells were chosen to study the Factor VII promoter in a hepatocyte-like environment. HeLa cells, derived from a human cervical carcinoma, were chosen as a nonhepatocyte cell line.

Optimization experiments were first performed to determine the quantity of DNA and incubation times to use in transient CaPO₄ transfections. For HepG2 cells, 3 μg of test plasmid DNA (1 pmol) in 12-h incubations were used; samples were taken 48 h later. Identical conditions were used for transfections of HeLa cells; however, transfections were performed with 15 μg of test plasmid DNA (5 pmol). To correct for variations in transfection efficiency in both HepG2 and HeLa cell experiments, a luciferase internal control plasmid (2.5 μg) was cotransfected along with the growth hormone reporter gene constructs. All results were normalized in relation to luciferase expression (Fig. 4A). Results in HepG2 cells showed that the first 75 bp alone were insufficient to increase reporter gene activity above base line. However, the addition of the adjacent 5′ 34 bp in a 109-bp construct dramatically increased promoter activity to approximately 20-fold above base line. Promoter activity increased slightly upon including the next 53 bp (162-bp construct) and again with adding the next 24 bp (186-bp construct). The 186-bp construct, −185 to +1, conferred maximal activity. A statistically significant decrease (p < 0.01) to 59% of the 186-bp construct was seen after including the adjacent 105 bp. Levels increased again at 501 bp and rose to a maximum at 1.746 kb. However, the increase above the promoter activity seen with the 186-bp construct was not significant. A decrease was seen between 1.7 and 3.9 kb.

HeLa cells were used to compare promoter activity to that seen in HepG2 cells in order to study differences between a cell with hepatocyte-like characteristics and a nonhepatocyte derived cell. The 109-bp construct had maximal promoter activity but showed levels only 3 times the base-line value. For the remaining constructs, including 162 bp or more, activity was at or below base line. These results indicate that the Factor VII promoter is more active in HepG2 cells than in HeLa cells.
The promoter activity of two different polymorphic alleles was compared in HepG2 cells; one construct contained the FVII sequence from an individual carrying a decanucleotide insert at -323 (bases -324 to -333); the other did not contain this insert (23). The construct containing the decanucleotide insert had a promoter activity that was only 67% that of the more common allele lacking the insert when assayed in four independent transfection experiments with duplicate preparations of each of the two constructs. This difference reached a statistical significance of p < 0.005 using the Student's t test to compare the results (67% ± 21% for the construct containing the decanucleotide insert and 100% ± 23% for the construct lacking the insert (n = 28)).

Identification of Protein Binding Sites in the Factor VII Promoter by DNase I Footprint Analysis and Gel Mobility Shift Assays—DNase I footprint analysis of the proximal 309 bp of 5'-flanking DNA of the FVII gene showed three protein-protected sites (Fig. 5): region A (-51 to -32), region B (-108 to -84), and region C (-233 to -215). These regions were protected using both HepG2 and HeLa nuclear extracts as shown in lanes 4 and 5. The protected DNA sequence in region B includes a G-rich sequence (lower strand) at -101 to -94, which contains a previously described Sp1 binding site (CCCCCTCCCCC) (44).

Gel mobility shift assays confirmed the presence of proteins, which bind specifically to regions A, B, and C. Fig. 6A shows the gel mobility shift assay using an oligonucleotide spanning the sequence between -51 and -29 bp. In HepG2 and HeLa nuclear extracts, two faint bands are seen (lanes 1 and 3) in addition to a faster migrating complex. A slower migrating band is also seen (top arrow) more prominently in HeLa than in HepG2 nuclear extracts. The ability of cold competitor to eliminate the binding of the radiolabeled oligonucleotide indicates the specificity of this sequence for the nuclear proteins.

Fig. 6, B and C, shows gel mobility shift assays using HepG2 nuclear extracts (lanes 1–4), HeLa nuclear extracts (lanes 5–8), and human liver nuclear extracts (lanes 9–12) with probes designed to study regions B and C as defined by DNase I footprint analysis (Fig. 5). Two slowly migrating complexes are seen using an oligonucleotide spanning the sequence -108 to -83 (Fig. 6B). Both complexes are seen more prominently using the HeLa nuclear extracts (lanes 5–8) than using either HepG2 (lanes 1–4) or liver nuclear extracts (lanes 9–12), suggesting a higher abundance of proteins in these complexes in HeLa cells. Both also disappear in the presence of cold competitor, suggesting that binding is specific for this DNA sequence.

A fast migrating complex is also present; this complex disappeared only in the presence of the highest concentrations of competitor (for HeLa and human liver nuclear extracts). In Fig. 6C, using a probe that extended from -236 to -201, two closely spaced complexes are seen that are reliably competed away by excess cold competitor (lanes 2–4, 6–8, and 10–12). A third faster migrating band was variably present and not thought to represent specific binding to this sequence.

An Sp1 Binding Site Is Identified in the -108 to -83 bp Region—Data from gel mobility shift assays using an oligonucleotide from -108 to -83 suggest that Sp1 is the protein present in the slower migrating complex (Fig. 7A). An Sp1 consensus oligonucleotide is able to compete away only the slower migrating band (lane 5), whereas cold self competitor is able to compete away both of the closely spaced complexes (lane 2). An Sp1 antibody, which does not cross-react with known Sp1-like transcription factors, supershifts the slower migrating complex (lane 3). Additionally, purified Sp1 binds to the oligonucleotide and migrates to the same position as the slower migrating band (lane 4). All of these data strongly support the presence of an Sp1 binding site within the -108 to -83 bp sequence.

Two mutations were made in the G-rich (lower strand) region between -103 and -92 base pairs at residues -100 and -94 by changing cytosine residues to adenosines. An oligonucleotide probe containing this -94/-100 mutation (MT94/100) fails to bind either the slower migrating protein or the protein in complex B (Fig. 7A, lane 6) and no longer binds to purified Sp1 protein (Fig. 7, lane 8). Independent mutations of the -94 and -100 bp sites revealed that the C to A change at -100 noticeably reduced binding on gel shift assay (Fig. 7B, lane 3) whereas mutation of the -94 bp site alone has a more profound effect, completely eliminating binding of both Sp1 and complex B on gel shift assays (Fig. 7B, lane 2). This result, the loss of
two distinct complexes on gel shift assay with the mutation of a single nucleotide, suggests the existence of overlapping binding sites for Sp1 and another nuclear protein.

HNF-4 Antiserum Supershifts Binding of Liver Nuclear Protein to a Factor VII Proximal Promoter Sequence—A 30-bp oligonucleotide sequence from the Factor VII promoter (−76 to −47), which contains an ACTTTG element at −63 to −58, was analyzed using gel mobility shift assays. The oligonucleotide was shown to bind a nuclear protein present in HepG2 cells and human liver but not in HeLa cells (Fig. 8A, lanes 1, 4, and 7). The binding could be competed away by incubating with excess unlabeled self competitor oligonucleotide (Fig. 8A, lanes 2 and 3 and lanes 5 and 6). Polyonal antiserum directed against the carboxyl terminus of HNF-4 supershifted the complex formation by liver nuclear extracts (Fig. 8B). Furthermore, mutation of the ACTTTG element to GACAAT eliminated binding to the liver nuclear extract (Fig. 8A, lane 9).

Mutagenesis of the Sp1 and HNF-4 Binding Sites Reveals Critical Elements in the Factor VII Promoter—Mutagenesis was done to analyze the functional consequences of mutations in the binding sites for Sp1 and HNF-4 (Fig. 4B). Mutations installed at −94 and −100 (MT94/100) in the Sp1 binding site and at −94 alone (MT94) reduced promoter activity to 2% of the wild-type 109 bp construct in HepG2 cells. In contrast, a mutation at −100 alone, which moderately affects binding on gel mobility shift assay, reduced promoter activity to 35% of wild-type.

To study the functional importance of this ACTTTG element, a 6-bp mutation (Δ ACTTTG: ACTTTG → GACAAT) was installed in the 186-bp pOGH construct and used in transfection experiments with HepG2 cells and in transactivation studies with HeLa cells. This 6-bp mutation (Δ ACTTTG) was previously shown to dramatically reduce activity in the Factor X promoter (7). In HepG2 cells, mutation of the putative HNF-4 binding site reduced promoter activity to 2% of the 186-bp wild-type (Fig. 4B). Transactivation experiments were performed to assess whether the presence of HNF-4 in trans could increase expression of the human growth hormone reporter gene above base-line levels. A 4-fold increase in promoter activity of the 186-bp wild-type construct was seen when cotransfected with 10μg of pLEN4S, an expression plasmid containing the HNF-4 cDNA. This was not seen with cotransfection of an equimolar amount of the pLEN expression plasmid, lacking the HNF-4 cDNA insert (wild-type + pLEN4S: 392% ± 33.9% versus wild-type + pLEN: 100% ± 9.6%). When this experiment was repeated using a 186-bp reporter gene construct

---

**Fig. 6.** Gel mobility shift assays confirm binding of nuclear factors to the protected regions (A, −51 to −32; B, −108 to −84; and C, −233 to −215) of the human Factor VII promoter as determined by DNase I footprinting. Areas protected in the DNase I footprint analysis (Fig. 5) were analyzed for their ability to bind specifically to nuclear extracts in three different cell types: HepG2 cells, HeLa cells, and human liver cells. Increasing concentrations of unlabeled self competitors are used to differentiate specific from nonspecific binding. A, an oligonucleotide probe containing the sequence between −51 and −29 shows formation of four complexes (indicated by arrows) with nuclear extracts from HeLa cells (lanes 1 and 2) and HepG2 cells (lanes 3–6). Lane 2 contains 100-fold of cold self competitor. Lanes 4–6 contains self competitors at 10-fold, 100-fold, and 500-fold excess, respectively. B, an oligonucleotide spanning the sequence from −108 to −83 was incubated with nuclear extracts from HepG2 (lanes 1–4), HeLa (lanes 5–8), and human liver (lanes 9–12). Specific DNA-protein complexes are indicated by arrows. Indicated amounts of molar excess of cold self competitors were added. C, an oligonucleotide spanning the sequence from −236 to −201 was incubated with nuclear extracts from HepG2 cells (lanes 1–4), HeLa cells (lanes 5–8), and human liver (lanes 9–12). Two specific complexes are indicated by thick arrows. A faster migrating band as indicated by a thin arrow is not consistently seen and is therefore considered nonspecific. Molar excess of cold self competitors were added as indicated. Nucleotide sequence of the probe is shown at the bottom.
DISCUSSION

A longstanding assumption in the field of coagulation has been that the vitamin K-dependent coagulation proteins are synthesized exclusively in the liver. Recent data indicate, however, that this assumption is not true, at least at the mRNA level, where several groups have documented the presence of transcripts in non-liver tissues of protein S (45), protein C (46), prothrombin (47), and Factor X. For human coagulation Factor VII in contrast, within the sensitivity of the assay, Northern analysis using poly(A)⁺ RNA from a wide variety of tissues demonstrates that expression is confined to the liver (Fig. 1). Thus, at least for Factor VII, the older notion in fact appears true.

Factor VII, along with Factor X, is located on the long arm of chromosome 13 (13q34) (48) only 2.8 kb 5' of the Factor X gene (8). Densitometric analysis of autoradiograms demonstrated that steady state levels of Factor VII mRNA were only 6% those of Factor X mRNA levels, correlating well with the 10–20-fold difference in plasma levels. Thus, the difference in plasma concentrations of the two factors is due to a difference in rates of transcription or in RNA stability (or some combination of the two).

A qualitative comparison of our Factor VII promoter data with those previously published for Factor X (7, 8) shows similarities in that both promoters are TATA-less with the essential regulatory elements located within the first several hundred bp 5' to the translation start site. An HNF-4 binding site, including the same ACTTGG motif seen in the promoters of Factors IX (9) and X (7, 8), is shown to be critical in the Factor VII promoter. Although a CAAT box, a conserved element present in approximately 30% of eukaryotic promoters (49), is critical in the promoters of both factor IX (9) and Factor X (7), no such element is apparent in the Factor VII promoter.

Identification of the transcription start site is an essential step in the characterization of a promoter. Like the promoters of many other clotting factor genes, the Factor VII promoter lacks a TATA box, a sequence present in approximately 80% of RNA polymerase II eukaryotic promoters (49), which is usually located about 30 bp upstream from the transcription start site. Many TATA-less promoters are thought to depend on the presence of the transcription factor Sp1 upstream from the start site (50, 51). TATA-less promoters are also frequently characterized by the presence of an initiator element encompassing the transcription start site (52). In this report, a major transcription start site is identified at −51 by both primer extension and RNase protection. This start site occurs in the context

complex. Incubation with 1 μl Sp1 antibody (lane 3) results in loss of this complex and appearance of a "supershifted" complex. Mutation of the cytosine residues at −108 to −104 to adenosines eliminates binding of both slowly migrating complexes from HeLa nuclear extracts (lane 6). Binding of purified Sp1 is seen with the wild-type oligonucleotide sequence (lane 4) but not with the mutated sequence (lane 8). B, mutation of cytosine to adenosine at position −94 but not at position −100 eliminates binding of Sp1 and the protein in complex B. Wild-type (lane 1) and mutant (lanes 2–4) oligonucleotides spanning −108 to −83 were incubated with HeLa nuclear extracts. Sequences of the probes are shown at the bottom. The specific DNA-protein complexes are indicated by arrows.

4 H.-L. Hung and K. A. High, unpublished observations.
TCA_{23}-GTCCC, a site that is in 100% agreement with both a mathematically calculated consensus sequence CANYYY (49) based on the cap signal of 502 eukaryotic promoters, and an experimentally defined consensus PpyPuPyNTA/Ppy determined using synthetically prepared promoter sequences (53). An Sp1 binding site, which we show to be crucial in promoter activity of Factor VII, is located approximately 41 base pairs upstream of the identified start site. This position corresponds well with the observed optimal distance between a GC-rich Sp1 binding element and the transcription start site (52, 54). It should also be noted that, as is the case for most other TATA-less promoters, multiple other start sites are identified in the region surrounding the major transcription start site.

Functional characterization of the human Factor VII promoter was performed in both HeLa cells, a non-hepatocyte-derived cell line, and HepG2 cells, a hepatoma-derived cell line previously shown to secrete Factor VII (55). Analysis of promoter activity using varying lengths of 5' flanking region adjacent to the translation start site revealed a dramatic increase in promoter activity from base line to maximal activity upon the addition of a G-rich (lower strand) region, which includes an Sp1 binding site. The importance of this binding site was demonstrated by a dramatically reduced promoter activity (2% of wild-type) in a construct containing a single base change (C to A), which eliminated binding of Sp1 by gel mobility shift assay. However, the possibility cannot be excluded that a second protein, with an overlapping binding site to that of Sp1, plays an equally important role.

In reporter gene assays, a reproducible and statistically significant decrease of 41% was seen between constructs containing 186 and 291 base pairs of Factor VII promoter sequence. This region includes a site at -233 to -215, which was protected on DNase I footprint analysis and bound nuclear proteins on gel mobility shift assays. Interestingly, part of the protected sequence (5'-PuGGTCA NTGACPCy-3') bears great resemblance to the sequence of several hormone response elements (56). Experimental studies testing the influence of spacing, orientation, and the exact sequence of the motif and flanking sequence have documented the importance of all these factors in determining receptor binding to a particular combination of half-sites (26). The sequence, which contains two PuGGTCA half-sites in an inverted orientation separated by 2 bp, does not exactly match any documented hormone responsive elements. It should be noted that putative binding sites for several other factors were also seen in this region.

Experiments revealed that an intact binding site for HNF-4, a known member of the steroid hormone receptor superfamily (35), is critical for Factor VII promoter activity in vitro. First, a protein present in both human liver and HepG2 nuclear extracts but not in HeLa cells could be supershifted with the addition of HNF-4 antiserum. Mutation of the sequence ACTTTG (−63 to −58), which is well documented to bind HNF-4 in the Factor IX promoter (9), eliminated binding in gel mobility shift assays. Furthermore, an HNF-4 expression plasmid was able to increase promoter activity 4-fold when cotransfected with a 186-bp Factor VII promoter construct in HeLa cells. Mutation of the ACTTTG not only dramatically reduced promoter activity in HepG2 cells (2% of wild-type) but also resulted in the loss of ability of the HNF-4 expression plasmid to transactivate the Factor VII construct in HeLa cells. These experiments demonstrate that the HNF-4 binding site is critically required for Factor VII promoter activity. Given these findings, it is perhaps surprising that protection of the HNF-4 binding site is not seen on footprint analysis, but in this case this appears to result from the resistance of this region of DNA to digestion by DNase I, since no ladder is seen in the "no

**Fig. 8.** A, gel mobility shift assays reveal binding of a protein from liver and HepG2 nuclear extracts but not HeLa nuclear extracts to an oligonucleotide sequence containing the ACTTTG motif (bases −63 to −58) in the proximal Factor VII promoter. A probe spanning the sequence from −76 to −47 forms a complex (arrow at left) with a protein present in HepG2 nuclear extracts (lanes 1–3) and human liver nuclear extracts (lanes 4–6). Addition of excess unlabeled self competitor results in a gradual disappearance of the complex at 10× (lanes 2 and 5) and 100× (lanes 3 and 6) concentrations. HeLa nuclear extracts do not show binding of a protein in this region (lanes 7 and 8). Mutation of the ACTTTG to GACAAT between −63 and −58, within the −76 to −47 sequence, eliminates binding of human liver nuclear extracts to the mutated radiolabeled probe (lane 9). B, anti-HNF-4 antibodies supershift the binding of human liver nuclear extracts to a probe containing the sequence between −76 to −47 bp. Lane 1 shows the formation of a complex (left arrow) by the −76 to −47 radiolabeled probe with human liver nuclear extract. Lane 2 shows the binding of the same components with the addition of 1 μl of antiserum raised against the carboxyl terminus of HNF-4. The position of a slower moving immune complex is indicated (right arrow).
A recent report documented a reduced level of Factor VII antigen in individuals heterozygous for a decanucleotide insert at position −323 in the Factor VII promoter (81% antigen level in individuals heterozygous for the insert, 112% antigen level in individuals without the insert, p < 0.005) (24). We investigated whether the presence of this decanucleotide in reporter gene constructs could influence promoter activity in HepG2 cells. Our findings showed that a 33% reduction in activity was associated with the naturally occurring allelic sequence containing the insert. This provides additional evidence that the presence of a decanucleotide insert may be directly related to the lower Factor VII levels seen in the population tested.

A schematic model of the promoter region of Factor VII based on data presented in this report is shown in Fig. 9. Four cis elements were shown to bind nuclear proteins. These elements are located at −51 to −32, −63 to −58, −108 to −84, and −233 to −215. Binding sites for the liver-enriched transcription factor, HNF-4, and the ubiquitous factor, Sp1 were identified at −63 to −58, and −101 to −94, respectively. A major transcription start site was identified at approximately −51, located between two protein binding regions. This start site was favorably located approximately 40 bp downstream of a functionally important Sp1 binding site. A sequence with homology to a Sp1 binding site was identified at approximately −51, located between two protein binding regions. This start site was favorably located approximately 40 bp downstream of a functionally important Sp1 binding site. A sequence with homology to a Sp1 binding site was identified at approximately −51, located between two protein binding regions. This start site was favorably located approximately 40 bp downstream of a functionally important Sp1 binding site.