Transcriptional Control of the Factor IX Gene: Analysis of Five Cis-Acting Elements and the Deleterious Effects of Naturally Occurring Hemophilia B Leyden Mutations

By David J. Picketts, Christopher R. Mueller, and David Lillicrap

Hemophilia B Leyden is a rare form of inherited factor IX deficiency in which patients experience spontaneous postpubertal recovery of factor IX levels. The mutations resulting in this disorder are localized in a 40-nucleotide region encompassing the major transcriptional start site for factor IX. Here we report the further characterization of five cis-acting elements in the factor IX promoter and the effects on protein binding and transcriptional activation of five Leyden mutations (at nucleotides +13, −5, −6, −20, and −26) that occur within the proximal three elements (sites 1 through 3). Bandshift studies using nuclear extracts from four different rat tissues have shown that at least some of the proteins binding to each of the five sites are ubiquitous in nature. The pattern of DNA binding at site 1 suggests that this element plays an important role in mediating the liver-specific expression of factor IX. Additional studies with liver nuclear extracts obtained at several different points in development have shown an increase in DNA binding at sites 1, 4, and 5 between 1 day and 1 week. Using DNase I footprint analysis and competition bandshift studies, we have shown that the binding of nuclear proteins to each of the mutant sites is disrupted to a variable extent. There appears to be some, although reduced, protein binding to all of the mutant oligonucleotides apart from the −26 mutant. In vitro transcription assays have shown that each of the mutations reduces the global proximal promoter activity by 40%. Two double mutant promoters did not show any additional downregulation in the in vitro transcription assay. In experiments designed to assess the relative transcriptional activity mediated from each of the five sites independently, we have tested artificial homopolymer promoters of each site in the in vitro transcription assay. These studies show that sites 4 and 5 are the strongest activators and that transactivation from site 5 is further enhanced by the albumin D site-binding protein. In summary, these investigations show deleterious effects of each of the Leyden mutations tested on the binding of trans-acting factors and also show disruption of transcriptional activation in a functional in vitro transcription assay. Our results also show that cis-acting elements 4 and 5 are the principal activators of this locus.

© 1994 by The American Society of Hematology.
In studies using nuclear protein extracts from a variety of tissues, we have documented the tissue-specific binding of proteins to all five cis-acting sequences. In addition, using liver nuclear extracts, we have examined the developmental changes in the pattern of gel-retarded complexes at each of the five sites. In studies of the mutant promoters, we have shown decreased binding of trans-acting factors to the mutant cis-acting sequences and have documented, in a quantifiable in vitro transcription assay, that all the mutations result in decreased transcriptional activity. We have also shown that sites 4 and 5 are the strongest activators at this locus and that transactivation from site 5 can be further enhanced by the transcription factor DBP.

MATERIALS AND METHODS

Plasmid constructs. The factor IX wild type (WT) and −5 mutant pUC plasmids were described previously.12 They encompass nucleotides −285 to +21 of the factor IX promoter. For in vitro transcription reactions, the WT- and −5-pUC plasmids were digested with HindIII and Sac I and purified on low melting-point agarose. These fragments were then cloned into the G-free/TATA vector derived from the Alu-7 plasmid described in Gorski et al.14 To prepare this vector, the sequence containing the −35 to +22 region of the albumin gene and the G-free cassette was amplified by polymerase chain reaction (PCR) and cloned into the Sac I and EcoRI sites of the pBS* vector (Stratagene, La Jolla, CA). This produces a vector with the pBS* multilinker from HindIII to Sac I, upstream of the albumin TATA sequence and the G-free cassette. Inclusion of the albumin TATA sequence in this construct was required after initial studies of the native TATA-less factor IX promoter showed products derived from multiple transcriptional start sites. The WT G-free construct was used as a template to construct the other Leyden mutants by site-specific PCR mutagenesis using the method of Jones and Howard15 with the modifications described below. The mutant oligonucleotides used were:

F9 A+13G: 5'-ACCACCTTCA-CAGTCTGC; AAAGTGTCAGACGTGGTG-5';
F9 A-5T: 5'-GGTACAACATTAATCGTCCIIT, TTGATTAGCAGGAACCATG-5';
F9 G-6C: 5'-GGTACAACTAATCCACCTTA; TTGATTAGGTGGCTTG-5';
F9 T-20A: 5'-AGCTCAGCTTCTACTTGGT; TGCACTGAAATCA-5';
F9 G-26C: 5'-AGCTCAGCTTCTACTTGGT; and
TGACAGATGGCAAAC-5'.

Fragments spanning the multilinker cloning site were amplified using the mutant primer and the universal reverse-sequencing primer to the factor IX promoter. For in vitro transcription reactions, the WT- and −5-pUC plasmids were digested with HindIII and Sac I and purified on low melting-point agarose. These fragments were then cloned into the G-free/TATA vector derived from the Alu-7 plasmid described in Gorski et al.14 To prepare this vector, the sequence containing the −35 to +22 region of the albumin gene and the G-free cassette was amplified by polymerase chain reaction (PCR) and cloned into the Sac I and EcoRI sites of the pBS* vector (Stratagene, La Jolla, CA). This produces a vector with the pBS* multilinker from HindIII to Sac I, upstream of the albumin TATA sequence and the G-free cassette. Inclusion of the albumin TATA sequence in this construct was required after initial studies of the native TATA-less factor IX promoter showed products derived from multiple transcriptional start sites. The WT G-free construct was used as a template to construct the other Leyden mutants by site-specific PCR mutagenesis using the method of Jones and Howard15 with the modifications described below. The mutant oligonucleotides used were:

F9 A+13G: 5'-ACCACCTTCA-CAGTCTGC; AAAGTGTCAGACGTGGTG-5';
F9 A-5T: 5'-GGTACAACATTAATCGTCCIIT, TTGATTAGCAGGAACCATG-5';
F9 G-6C: 5'-GGTACAACTAATCCACCTTA; TTGATTAGGTGGCTTG-5';
F9 T-20A: 5'-AGCTCAGCTTCTACTTGGT; TGCACTGAAATCA-5';
F9 G-26C: 5'-AGCTCAGCTTCTACTTGGT; and
TGACAGATGGCAAAC-5'.

RESULTS

The occurrence of natural mutations in the factor IX promoter, which result in a similar developmentally regulated phenotype, provides a unique opportunity to study the function of this clinically significant promoter. We now present new information relating to the five cis-acting elements in the promoter by examining the tissue-specific and developmental regulation of the factors interacting with these sites and have assessed the effect of five different Leyden mutations on the binding of these proteins and on in vitro transcription.

Tissue and developmental bandshift assays. Previous reports assessing the pattern of trans-acting proteins binding to the factor IX promoter have used a variety of purified and recombinant protein preparations as well as liver nuclear extracts. In this study, we have used bandshift assays employing nuclear extracts from four different tissues (liver, spleen, kidney, and brain) and a human hepatoma cell line (HepG2) to assess the tissue specificity of the factors binding to all five of the cis-acting sites (Fig 1A). A very heterogeneous complex is formed between site 1 and factors present in the liver, with no discernible difference between extracts prepared in the morning and at night. Several more distinct

From www.bloodjournal.org by guest on September 24, 2017. For personal use only.
complexes are formed with factors present in other tissues and in the dedifferentiated HepG2 extracts. In each case, the complexes are significantly less abundant than the factors in the liver (fivefold difference). The pattern of a DNA-protein "smear" seen with the liver extracts at site 1 is typical of a site that interacts with members of the C/EBP family and represents binding by a heterogeneous mixture of homodimers and heterodimers of these factors. This is consistent
FACTOR IX GENE TRANSCRIPTION

with the previous characterization of this site as a high-affinity C/EBPα site,6,12 a finding that is further substantiated by the results of a supershift assay for C/EBPα at this site (Fig 1C). This analysis also indicates that site 1 binds C/EBPα and DBP to a lesser extent.

Sites 2 and 3 are characterized by what appears to be a discrete complex present in all of the extracts. Further analysis of these complexes by fractionation of the proteins on heparin agarose indicates that the site 2 protein is likely a single complex that elutes at 400 mmol/L NaCl. In contrast, the factors binding to site 3 can be separated into two distinct fractions eluting at 200 and 500 mmol/L NaCl. This observation is consistent with the fact that site 3 has previously been characterized as binding both HNF-4 and the androgen receptor.11,13

Site 4 is known to bind NF 1, which is present in all tissues, and the more abundant complex formed in the liver is caused by the presence of more liver-specific forms of this factor.14 Experiments with recombinant C/EBPα indicate that this factor binds to this site to a limited degree.12 Site 5, in our previous studies, was shown to bind an unknown factor in addition to C/EBPα. This unknown factor, which forms a discrete complex, appears to be present in all of the other tissues examined, suggesting that it may be a ubiquitous protein.

Examining the developmental regulation of all of these sites in the liver indicates that a significant induction of DNA binding activity (threefold increase) occurs at sites 1, 4, and 5 between day 1 and 1 to 2 weeks of age (Fig 1B). This is the period when terminal differentiation occurs and is associated with the induction of factors such as C/EBPβ, which binds to site 1. A second complex with site 3 was also observed from 1 to 2 weeks of age. This is before puberty in the rat and, thus, is unlikely to represent induction of the androgen receptor. No specific changes are observed between 1 month of age (peri-pubertal) and the adult, which corresponds to the period in humans when phenotypic recovery occurs. However, given the difficulty of resolving multiple complexes at a given site, this finding is not unexpected.

Effect of Leyden mutations on protein binding. We recently reported that an A to T transversion at nucleotide −58 greatly decreased the affinity for protein binding to site 2 of the factor IX promoter in DNase I protection assays.15 To determine if other Leyden mutations at each of the three proximal sites disrupted protein binding to the corresponding mutant sites and to assess the possible disruption of protein binding to adjacent sites, we synthesized promoter templates corresponding to mutations at +13 (A to G; site 1), −6 (G to C; site 2), −20 (T to A; site 3), and −26 (G to C; site 3). Each of these templates was studied in DNase I footprint assays using rat-liver nuclear extracts. All of the mutations, with the exception of the substitution at +13, resulted in the loss of DNase I protection at the site in which the mutation was located (Fig 2). The +13 mutation located in site 1 has previously been shown to disrupt the binding of C/EBPα.9 Repeating the footprints of the +13 mutant template with both heat-treated extracts, which preferentially retain C/EBPα activity, and with recombinant C/EBPα showed that C/EBPα could not occupy site 1 when the mutation was present (data not shown). The footprint obtained with non-heat-treated extracts is presumably the result of other factors such as C/EBPβ binding to this site.

Oligonucleotides for sites 1, 2, 3, and each of the Leyden mutant constructs were radiolabeled and examined in gel mobility shift assays for protein binding (Fig 3, A through C). The +13 and −5 mutant probes showed slightly reduced complex formation compared with the WT probes (70% and 86% binding, respectively). In contrast, the −6, −20, and −26 mutant oligonucleotides showed very poor complex formation (5%; 3%; and 0%, respectively). To assess the relative affinities of the mutant sites, protein binding to radiolabeled WT oligonucleotides was competed with either WT or mutant unlabeled oligonucleotides at 50 or 100 times the WT probe concentration. The +13 oligonucleotide showed poor competition, with a discrete retarded complex remaining even with a 100-fold excess of the unlabeled mutant competitor. This presumably represents the preferential competition of all of the other site-1 binding factors with the exception of C/EBPα, which forms the remaining complex. The −5 oligonucleotide competed poorly for the site-2 complex and the −6 oligonucleotide showed no competition, indicating that the −6 mutation has a more profound effect on binding of the currently unknown protein that interacts with site 2. Studies with the −20 and −26 oligonucleotides showed markedly reduced competition compared with the WT site-3 probe. Two other groups have shown that oligonucleotides containing the −20 or −26 nucleotide change compete weakly (the −20 mutation) or not at all (the −26 mutation) for HNF-4 binding to this site.11,13

In vitro transcription analysis of full-length mutant promoters. Having documented the disruption of protein binding at each of the mutant promoter sites in footprint and bandshift experiments, we have gone on to perform studies to assess the functional consequences of this effect. An in vitro transcription assay was used18 in which the WT and mutant promoters were cloned upstream of a G-free vector that contained the −35 to +22 region of the albumin gene including the TATA box (G-Free/TATA cassette).19 As described above, the inclusion of the TATA sequence in this construct was required to produce a single discrete transcript that could be readily quantified. The activity of the different mutants was compared with the transcriptional activity of a G-free cassette under the control of the adenovirus major late promoter. Although this assay may have limitations in exactly reflecting the magnitude of the in vivo effect of these mutations, it is an excellent in vitro method for comparing the relative effect on transactivation of each of the five promoter mutants. Figure 4 shows the results of a single experiment in panel A and the mean values obtained from five experiments in panel B. All of the mutations resulted in a decrease in transcriptional activity to 60% or less of the WT. An additional examination of the activity of two double mutants (+13 and −5 [sites 1 and 2], and −5 and −20 [sites 2 and 3]) that retain the function of sites 3 and 1, respectively, showed that they were not significantly more debilitated than the lowest of the two single mutations in each combination (Fig 4B).

The extent of transcriptional disruption resulting from the
Leyden mutations in this assay, compared with their in vivo effects, suggests that other regulatory elements that are not being tested in this system may also play a role in generating the Leyden phenotype.

**In vitro transcription analysis of homopolymeric elements.**

To assess the independent transactivation potential of each of the five *cis*-acting sequences in the proximal promoter, oligonucleotides corresponding to each of the five sites were concatamerized and cloned into the G-freeRATA cassette. This strategy has been shown previously to be an effective method to characterize the role of individual *cis*-acting elements in a complex promoter.¹⁶⁻²⁰

The number of monomeric sites contained in each of the constructs ranged from 5 to 11. In the case of the albumin promoter, previous studies have indicated that no significant difference in transcriptional activation is evident in homopolymers containing between 5 and 12 copies of the monomeric site (unpublished data). The homopolymeric constructs were examined for promoter activity, standardized with the adenovirus major late-promoter activity and compared with the activity of the G-Free/TATA plasmid that did not contain an insert (Fig 5, A and B). Sites 4 and 5 were the strongest transactivators with site 1 also showing some activity. The liver specificity of site-5–mediated transactivation is also supported by the observation that an in vitro transcription assay performed with kidney nuclear extracts shows only ~18% of the activity seen with liver extracts (data not shown). In this assay, sites 2 and 3 do not appear to mediate independent transcriptional activation. These results suggest that transactivation of the factor IX promoter is mainly caused by the binding of factors to sites 4 and 5, and that the three proximal sites may play a different role in mediating transcription from this locus.

Finally, in light of our previous observation concerning a transcriptional recovery mechanism involving the peripuberal onset of expression of DBP, we studied the effect of this trans-acting factor on each of the five sites in the in vitro transcription assay. Recombinant DBP was added to the in vitro transcription reactions of the homopolymer promoter sites and transcriptional activation was quantified with reference to the adenovirus control transcript (Fig 5, A and B). Only site 5 showed an increase in transcriptional activity with the addition of DBP.

**DISCUSSION**

In this study we have used a quantifiable in vitro transcription assay to examine the effects of various Leyden mutations on binding of cognate factors and on the efficiency of
transcription of this gene. This assay uses rat-liver nuclear extracts as a source of trans-acting factors to regulate transcription, and as such, may represent a more physiologically relevant test of transcriptional activation than other models based on cultured cells that may be deficient in critical trans-acting factors. First, the basis for the liver-specific expression of factor IX was examined using bandshift assays with nuclear extracts from four different tissues in conjunction with oligonucleotides corresponding to all of the factor IX elements. Only site 1 exhibited significantly more protein binding to it in the liver extracts. This is in keeping with the previous identification of site 1 as a high-affinity C/EBPα binding site, a finding that we have confirmed in a supershift assay in this study. Site 5 has also been identified as a C/EBPα binding site but in this study, the tissue-specific differences in C/EBP binding are masked by the presence of a non–tissue-specific protein whose identity is unknown. Thus, from the results of this study and our previous report, the primary tissue-specific elements for this promoter appear to be sites 1 and 5, with C/EBPα being the main tissue-specific regulator of this gene through these sites. Consistent with this observation is the strong transcriptional activity of
children. However, we have observed no apparent changes in the pattern of DNA binding activity at any of the five sites between the peripubertal period and adulthood, corresponding to the time when phenotypic recovery occurs in hemophilia B Leyden.

Our previous report of a potential mechanism for the spontaneous postpubertal improvement of the Leyden phenotype, suggests that the liver-enriched transcription factor, DBP, plays an important role in mediating this recovery.12 In the rat, the onset of DBP expression is in the peripubertal period, and DBP levels are higher in older rats.19 DBP appears to mediate its effect through an enhancement of binding of C/EBPα to the two distal promoter elements, with an especially marked effect at site 5.12 This effect may be masked in the bandshift assays by the presence of other factors; however, the addition of DBP to the in vitro transcription assay clearly results in significant transcriptional enhancement solely from site 5. In the context of the full-length promoter, we can

site 5 and, to a lesser extent, site 1. Site 4, which binds a liver-specific form of NF 1 as well as having a low affinity for C/EBPα, also shows some activator function. Thus, it appears that this site also contributes to a modest degree to the overall tissue-specific activity of the factor IX promoter.

Levels of circulating factor IX show a changing developmental profile in normal subjects. Normal childhood levels of factor IX are about 75% of adult levels.21 These levels increase by ≈25% into adulthood and continue to increase by ≈1% per year throughout life.21,22 C/EBPα, which appears to be the prime regulator of this gene, is expressed in the fetal liver and its level appears to stay relatively constant after this point.24 This likely establishes the basal transcriptional activity of the factor IX gene. The analysis of developmentliver nuclear extracts has shown a significant increase in DNA binding activity between 1 day and 1 week, at sites 1, 4, and 5. Most of the increase at site 1 is caused by the induction of C/EBPβ, which occurs at this time and is known to bind to this site. Some of the increase in site 5 activity is also caused by C/EBPβ, but in addition, appears to involve the induction of the unknown ubiquitous factor that also binds to this site. These changes may be responsible for the gradual increase in levels of factor IX that occur in normal

Fig 4. (A) Autoradiogram showing the results of a single in vitro transcription assay. The studies were performed with a WT factor IX promoter (WT), five mutant promoters (+13, -5, -6, -20, and -26), two double mutant promoters (+13/-5 and -5/-20), a promoterless G-free cassette, and the albumin promoter using rat-liver nuclear extracts. In each case, the experimental transcript was quantified after normalization with the corresponding transcript directed by the control of the AdML promoter. (B) Tabulated results of five in vitro transcription assays using the WT and various Leyden mutant promoters. The average transcriptional activity and standard deviation are detailed for each construct.

Fig 5. (A) An in vitro transcription assay with rat-liver nuclear extracts using homopolymers of the five cis-acting sequences in the factor IX promoter and a promoterless G-free cassette (-). The results on the right of the autoradiogram have been performed with recombinant DBP added to the liver nuclear extracts. In each instance, the experimental transcript has been quantified in relation to the corresponding control transcript directed by the AdML promoter. (B) Results of five in vitro transcription assays using the homopolymerized sites 1 through 5 of the factor IX promoter with (n = 3) and without (n = 6) added DBP.
postulate that the increased activity mediated from site 5 would likely lead to the general activation of transcription. This observation further supports the importance of this site in the recovery of the mutant Leyden promoters and may also play a role in the continued increase in factor IX levels throughout adulthood seen in normal subjects.

We have shown that all five of the Leyden mutations that we have studied result in decreased protein binding by both DNase I protection assays and competition bandshift assays. Using a quantifiable in vitro transcription assay we have shown that these same mutations decrease the efficiency of transcription of this gene by at least 40%. The ability of the +13 mutation to interfere with transcription, despite allowing other factors to still bind to site 1, indicates that C/EBPα is the primary effector for this site. Both the −5 and −6 mutations have similar effects on in vitro transcription despite the −6 mutation being much more delebitated in its binding of the site-2 factor. The situation with the mutations in site 3 is more complicated. As expected, the −20 mutation still binds some protein, presumably the androgen receptor, whereas the −26 mutation looses all binding activity. However, despite their differential binding capabilities, they are equally defective in promoter activity. Double mutants, which incorporate two different mutations, are no more defective than the single mutants.

In light of the peripubertal timing of recovery of the Leyden phenotype, the role of androgens in this disorder has been explored by several previous studies.13,25,36 The results of our experiments do not exclude an effect of the androgen receptor, which may further increase transcription from the mutant promoters. However, we propose that this effect may not be through a direct enhancement of transactivation, but rather by increasing protein binding to site 3, thereby improving the stability of the transcriptional initiation complex. A similar role has been proposed for the progesterone receptor, for which experimental evidence has been obtained showing that binding to a progesterone response element facilitates the formation of a stable preinitiation complex, thus augmenting transcriptional initiation.27 In further support of this proposal, site 3 was not an activator of transcription with nuclear extracts in our homopolymer in vitro transcription assay. Factor IX is a TATA-less gene26,29 and details of the cis-acting elements that optimize the assembly and stability of the general transcriptional apparatus in this situation are lacking. A consensus initiator element30 does not appear to be present, and one possibility is that elements 1 through 3 act together as a site of assembly of the preinitiation complex. Understanding the role that androgens play in controlling the expression of factor IX is further complicated by the fact that no one has documented strong transactivation from the native promoter after cotransfection of an androgen-receptor expression plasmid,13 and finally, if the androgen receptor had a strong effect on transcription of the factor IX gene, one might expect to see lower levels of factor IX in normal prepubertal males and differences in factor IX activity between males and females, neither of which is the case.32

In conclusion, our analysis of the independent contributions of each of the five cis-acting elements in the factor IX promoter shows that the two distal elements, sites 4 and 5, are the strongest mediators of transcriptional activation, whereas the three proximal elements, within which the Leyden mutations have been described, appear to act by another mechanism. Postpubertal recovery of the Leyden phenotype may be mediated by the combination of increased activation from site 5 because of the onset of DBP expression and binding at site 3 of the androgen receptor (Fig 6). Our data from the homopolymer in vitro transcription assays suggests that the role of androgen receptor binding at this site may not be to directly transactivate this locus.

REFERENCES

of a binding site for hepatocyte nuclear factor 4 results in hemophilia B Leyden. Proc Natl Acad Sci USA 89:6300, 1992


Transcriptional control of the factor IX gene: analysis of five cis-acting elements and the deleterious effects of naturally occurring hemophilia B Leyden mutations

DJ Picketts, CR Mueller and D Lillicrap