Localization of transcription initiation sites in the human coagulation factor IX gene

Marleen J. Reijnen, Rogier M. Bertina and Pieter H. Reitsma

Hemostasis and Thrombosis Research Unit, Department of Hematology, University Medical Center Leiden, Leiden, The Netherlands

Received 23 July 1990

The transcription start sites of the human gene coding for the coagulation factor IX have been identified. Three major transcription initiation sites within a small area of ~30 nucleotides were found by S1 nuclease analysis and primer extension studies.

Blood coagulation; Factor IX; Transcription initiation site

1. INTRODUCTION

Human coagulation factor IX (FIX) is a vitamin K-dependent glycoprotein that plays a key role in blood coagulation. It is a component of the middle phase of the intrinsic clotting pathway. In the activated form it interacts with factor VIIIa, phospholipids and calcium ions to form a complex that converts factor X to Xa [1].

The human FIX is encoded by a single functional gene which is located on the X-chromosome [2,3]. It contains 8 exons and 7 introns which cover a region of ~34 kb [4]. The gene is expressed primarily in the liver [5]. Using primer extension and S1 nuclease mapping Anson et al. [4] detected 3 start sites of mRNA synthesis (Fig. 1). Two of these were considered to be an artefact while the transcription initiation site at an A (Fig. 1, position +1) was pinpointed as the major start site. In contrast Salier et al. [6] recently claimed a major start site at position −150. No transcription initiation was observed at the previous proposed site at position +1.

The discrepancy between both observations has major implications for the interpretation of past and future studies of the FIX gene promoter. For instance, mutations at positions −20, −6 and +13 [7–9] have been found in hemophilia B patients with an altered promoter function (hemophilia B Leyden). Furthermore, mutation of position +13 results in the loss of liver specific transcription factor C/EBP binding [10]. Therefore, we re-examined the site of transcription initiation of the FIX gene. Using primer extension experiments and S1 nuclease mapping experiments on human liver RNA we identified 3 major transcription start sites at positions +1, +4 and +30. No transcription start site was found at position −150.

Correspondence address: M.J. Reijnen, Hemostasis and Thrombosis Research Unit, Department of Hematology, Bldg 1: C2-R, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands

2. MATERIALS AND METHODS

2.1. *Materials*

Reverse transcriptase and S1 nuclease were obtained from Pharmacia LKB Biotechnology Inc. and Bethesda Research Laboratories. [γ-^{32}P]dATP (>3000 Ci/mmol), [α-^{35}S]dATP (>600 Ci/mmol) and T4-polynucleotide kinase were purchased from Amersham International (Amersham, UK). The cDNA synthesis kit and RNasin were obtained from Promega Biotec (Madison, USA). Oligonucleotide primers were made on a Cyclone DNA synthesizer (Millipore, Bedford).

2.2. *RNA preparation*

Total RNA was extracted from human liver tissue, human lymphocytes and HL60 cells as described previously [11,12].

2.3. *Primer extension*

Primer extension was performed using 4 different oligonucleotide primers. The sequences in parentheses indicate the position of the primers in the complementary strand (numbering as in [13]).

120 (5′-$^{-82}$AAAGACCCATTGAGGGAGAT^{-63}-3′)

182 (5′-$^{+88}$TAGGATATCTACTCAGTCGT^{+107}-3′)

122 (5′-$^{+108}$GAATGTACAGTTTTTCTTGA^{+6335}-3′)

213 (5′-$^{+6373}$AGGTATAATTCAGGTAAATTGG^{+6394}-3′)

90 ng of oligonucleotide primer (120/182/122/213) was labeled by T4-polynucleotide kinase and [γ-^{32}P]dATP. One tenth of the reaction was annealed with 50 μg total RNA for ~16 h at 28°C in a 20 μl mixture containing 40 mM Pipes (pH 6.4), 0.4 M NaCl, 1 mM EDTA (pH 8), 80% formamide.

After hybridization the nucleic acids were recovered by ethanol precipitation. Thirty μl of a mixture containing 10 × first strand buffer (cDNA synthesis kit), 10 mM DTT, 1 mM each cold dNTP, 4 mM sodium pyrophosphate, 20 U RNasin and 10 U reverse transcriptase was then added. After 1 h incubation at 42°C, the reaction was stopped by adding 1 μl 0.5 M EDTA. The extended products were recovered by ethanol precipitation and suspended in 10 μl

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

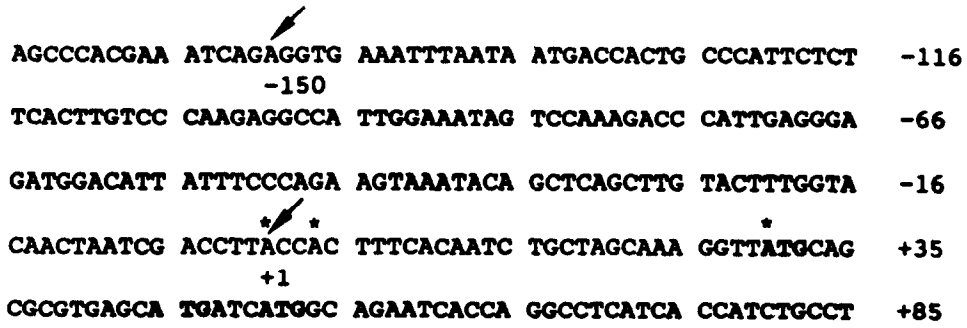

Fig. 1. Nucleotide sequence of part of the human coagulation FIX gene locus (numbering as in [7]). The transcription initiation sites found in this study are indicated (*). Arrows indicate the position of the transcription initiation sites proposed by Anson et al. [4] and Salier et al. [6]. Potential initiation codons are in bold type.

formamide-dye solution (80% formamide, 10 mM EDTA (pH 8.0), 1 mg/ml xylene cyanol FF and 1 mg/ml bromophenol blue). The mixture was then boiled for 5 min and loaded onto a 8% denaturing polyacrylamide gel. The size of each extended product was compared with the dideoxynucleotide derived sequencing ladder obtained with the same primer.

2.4. *S1 nuclease analysis*

S1 nuclease analysis with human liver RNA, human lymphocyte RNA and HL60 cell RNA was performed essentially as described by Maniatis et al. [12]. The ^{32}P-labeled fragments complementary to the mRNA were generated in a PCR reaction with 5′-end labeled antisense primers and unlabeled sense primers. S1 nuclease mapping experiments were performed using two different sets of oligonucleotide primers. To generate probe A we employed the following oligonucleotides in our PCR amplification: primer 117 (5′-$^{-81}$AAGACCCATTGAGGGAGATG^{-62}-3′) and primer 182 (see section 2.3: 5′-$^{+88}$TAGGATATCTACTCAGTGCT^{+107}-3′).

To generate probe B we used the primers 203 (5′-$^{-193}$CAGACTCAAATCAGCCACAG^{-174}-3′) and 120 (see section 2.3: 5′-$^{-82}$AAAGACCCATTGAGGGAGAT^{-63}-3′).

About 1–2 × 10^3 cpm of probe was dried down with 100 μg total RNA. The pellets were resuspended in a 20 μl mixture containing 80% formamide, 40 mM Pipes (pH 6.4), 1 mM EDTA (pH 8.0), 0.4 M NaCl. Samples were heated for 15 min at 85°C; then hybridization was performed for 16 h at 49°C. Three-hundred μl S1 nuclease buffer (0.28 M NaCl, 0.05 M sodium acetate (pH 4.5), 4.5 mM $ZnSO_4$, 20 μg/ml carrier ss DNA) were added to each sample and the reaction was started by the addition of 1000 U/ml S1 nuclease. After 30 min at 37°C, the reaction was stopped by adding 50 μl termination mix (4 M ammonium acetate and 100 mM EDTA).

The RNA-DNA hybrids were recovered by phenol/chloroform extraction followed by ethanol precipitation and were analyzed on an 8% acrylamide-7 M urea sequencing gel. The size of each protected fragment was compared with the dideoxynucleotide derived sequencing ladder obtained with the same primer (182 or 120).

3. RESULTS AND DISCUSSION

To determine the site of transcription initiation of factor IX mRNA, both primer extension and S1 nuclease mapping were used. The 5′-ends of FIX mRNAs were mapped by S1 protection of liver RNA, lymphocyte RNA and HL60 RNA. Each RNA sample was assayed with a S1 probe derived from the −81 to +108 region of FIX gene (probe A). The strategy for the mapping experiment is shown in Fig. 2A. The results with probe A show a heterogeneous pattern of initiation, with 3 clusters of 5′-ends in a 30 base region situated at position +1 to +31 (Fig. 2B). This complex of 5′-ends was absent when lymphocyte RNA or HL60 RNA were used. Comparison with the dideoxynucleotide derived sequencing ladder indicated that in the human FIX gene the 3 major initiation sites are located at positions +1, +4 and +30. A minor single band was observed at position +8.

To confirm that the S1 nuclease protected fragments represent transcription initiation sites, primer extension experiments were performed. Synthetic oligonucleotides 182 (exon I), 122 (exon I–exon II splice) and 213 (exon II, not shown) were hybridized to RNA from human liver, human lymphocytes and HL60 cells (Fig. 3A). Following reverse transcription and analysis on a denaturing gel, several major extended products were observed whose lengths corresponded to the transcription initiation sites +1, +4 and +30 as determined by S1 nuclease analysis (Fig. 3B). One additional stronger band at position +68 was found. This band, which does not correspond to the results of the S1 nuclease analysis, is probably an artefact. Finally a minor start site at position +8 was detected in both S1 nuclease mapping and primer extension analysis (Figs 2 and 3).

Our results indicate that major transcription initiation sites are situated at positions +1, +4 and +30 in the human FIX gene. The locations of the transcriptional start site clusters are comparable to those found previously [4]. However, there is a difference in the interpretation of the results. Anson et al. pinpointed only the start site at position +1 as the major start site, whereas the other two start sites of mRNA synthesis were interpreted to be artefacts due to mRNA nicking. Since these two bands were also detected in our system both by S1 nuclease and primer extension experiments we conclude that they represent two true additional mRNA start points.

We presume that the heterogeneity of the transcription initiation sites in the FIX gene is caused by the

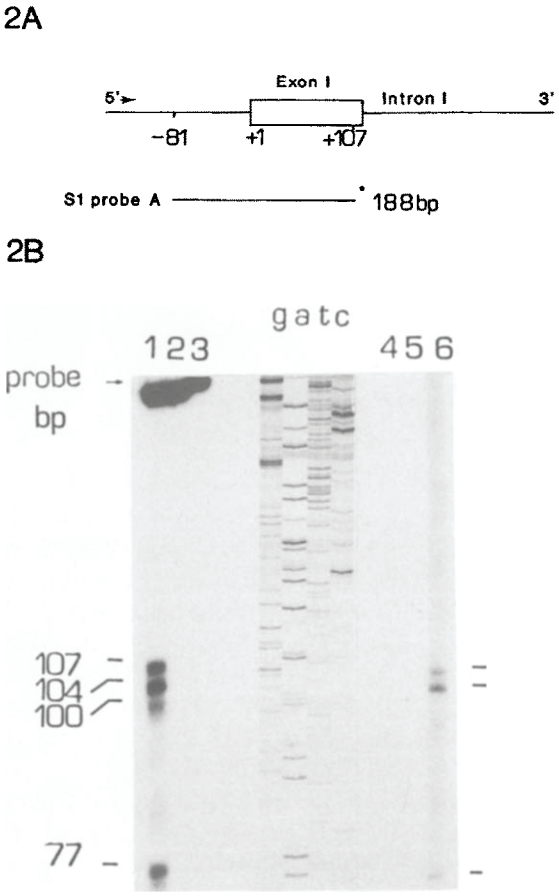

Fig. 2. Mapping the 5'-ends of human FIX mRNA. The segment of DNA used as probe (probe A) for S1 nuclease mapping is indicated (A). Primer extension analysis with oligonucleotide primer 182 is shown on the right. The same primer was used to generate probe A for S1 nuclease mapping as described in section 2. S1 resistant fragments and primer extension products were analyzed on an 8% acrylamide-7 M urea gel. The DNA sequence shown in (B) was obtained with the primer 182. S1 nuclease mapping: lane 1, total human liver RNA; lane 2, HL60 RNA; lane 3, total human lymphocyte RNA; dideoxy GATC track. Primer extension reaction: lane 4, total human lymphocyte RNA; lane 5, total HL60 RNA; lane 6, total human liver RNA. The length of protected fragments is indicated in bp.

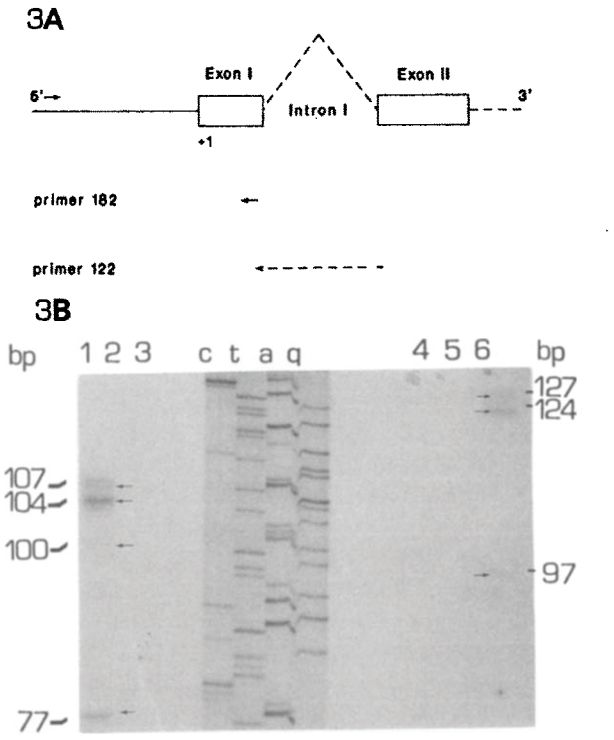

Fig. 3. Identification of the transcriptional starts of the human FIX gene with two different primers. Oligonucleotide primers 182 and 122, as indicated in Fig. 3A, were hybridized to RNA from human liver, lymphocytes and HL60 cells. The DNA sequence shown in (B) was obtained by priming a DNA clone of a human factor IX minigene in which the factor IX promoter is ligated to a full length cDNA (a gift from C. Schoemaker) with primer 182. Primer extension with primer 182: lane 1, liver RNA; lane 2, HL60 RNA; lane 3, lymphocyte RNA; dideoxy CTAG track. Primer extension with primer 122: lane 4, lymphocyte RNA; lane 5, HL60 RNA; lane 6, liver RNA. The length of the extension products is indicated in bp.

absence of a traditional TATA box as observed in many other TATA box lacking genes [14].

Experiments with a FIX chloramphenicol acetyltransferase (CAT) chimeric gene suggested a major transcription initiation site at −150 [6]. To confirm this transcription start site in vivo we have performed additional S1 nuclease experiments with probe B (described in section 2). No protected fragments were found in the upstream region between −193 and −63 (results not shown). To confirm this result a primer extension reaction was performed. For this purpose the synthetic primer 120 (−82 to −63) was used to specifically initiate reverse transcription of mRNA that starts at position −150 (results not shown). This primer was also used to initiate Sanger's dideoxy sequencing reaction. No extension products were found in this region, indicating that this site is more likely to be an artefact of the FIX-CAT chimeric system rather than a true transcription initiation site in vivo.

In conclusion, our experiments identified in human liver 3 major transcription start sites at positions +1, +4 and +30. This finding confirms the notion that the mutations in hemophilia B Leyden, that disrupt promoter function and C/EBP binding, are close to the transcriptional starts [7–10]. In future promoter studies it may be best to take the transcription start sites into account and include at least the +1 to +31 region in expression vectors.

REFERENCES

[1] Thompson, A.R. (1986) Blood 67, 565–572.
[2] Boyd, Y., Buckle, V.J., Munro, E.A., Choo, K.H., Migeon, B.R. and Craig, I.W. (1984) Ann. Hum. Genet. 48, 145–152.
[3] Camerino, G., Grzeschik, K.H., Jaye, M., De La Sall, H., Tolstosher, P., Lecocq, J.P., Heilig, R. and Mandel, J.L. (1984) Proc. Natl. Acad. Sci. USA 81, 498–502.

[4] Anson, D.S., Choo, K.H., Rees, D.J.G., Giannelli, F., Gould, J.A., Huddleston, J.A. and Brownlee, G.G. (1984) EMBO J. 3, 1053–1060.
[5] Fair, D.S. and Bahnak, B.R. (1984) Blood 64, 194–204.
[6] Salier, J.P., Hirosawa, S. and Kurachi, K. (1990) J. Biol. Chem. 265, 7062–7068.
[7] Reitsma, P.H., Bertina, R.M., Ploos van Amstel, J.K., Riemens, A. and Briet, E. (1988) Blood 72, 1074–1076.
[8] Kurachi, K., Hirosawa, S., Fahner, J.B., Wu, C.T. and Salier, J.P. (1989) Thromb. Haemostas. 62, 154.
[9] Reitsma, P.H., Mandalaki, T., Kasper, C.K., Bertina, R.M. and Briet, E. (1989) Blood 73, 743–746.
[10] Crossley, M. and Brownlee, G.G. (1990) Nature 345, 444–446.
[11] Affray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303–314.
[12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[13] Yoshitake, S., Schach, B.G., Foster, D.C., Davie, E.W. and Kurachi, K. (1985) Biochemistry 24, 3736–3750.
[14] Reynolds, G.A., Goldstein, J.L. and Brown, M.S. (1985) J. Biol. Chem. 260, 10369–10377.