RESEARCH ARTICLE

Single-Strand Conformation Polymorphism (SSCP) Analysis of the Molecular Pathology of Hemophilia B

Dezso David*, Humberto A.V. Rosa, Susan Pemberton, Maria J. Diniz, Manuel Campos, and João Lavinha

Laboratório de Genética Humana, Instituto Nacional de Saúde Dr Ricardo Jorge, Lisboa (D.D., H.A.V.R., J.L.); Haemostasis Research Group, Clinical Research Centre, Harrow, United Kingdom (S.P.); Serviço de Imuno-Hemoterapia, Hospital S. José, Lisboa (M.J.D.); Serviço de Hematologia, Hospital Geral de Santo António, Porto, Portugal (M.C.); Fax: 1-759-0441

Communicated by Edward Tuddenham

In the present study, we report the application of polymerase chain reaction-single-strand conformation polymorphism (PCR-SSCP) analysis to the screening of seven functionally important factor IX gene (FIX) regions (total length 2.66 kb) in 9 unrelated haemophilia B patients of Portuguese or African origin. In eight of the patients an altered migration pattern of single-stranded DNA was observed. Direct sequencing of the relevant DNA fragments unveiled the following sequence alterations: two novel mutations, namely FIX_{Barcelos} Thr-380-Pro and FIX_{Louisa} 9bp insertion at position 31,309 or 31,318; five mutations previously reported in other ethnic groups (FIX_{Porto} Arg-145-His, FIX_{Luanda} Gly-207-Arg, FIX_{Penafiel} Arg-248-Gln, FIX_{Sesimbra} Arg-333-Gln, FIX_{Cascais} Arg-333-Stop); and a normal variant, G → T transversion at position 6,596 in intron 2. We propose hypothetical models for the generation of the 9 bp duplication (FIX_{Louisa}). We have performed molecular modeling studies in order to predict the structure of the variant FIX molecules. © 1993 Wiley-Liss, Inc.

KEY WORDS: PCR-SSCP, FIX regions, Hemophilia B mutations

INTRODUCTION

Haemophilia B is an X-linked bleeding disorder caused by mutations at the factor IX (FIX) gene, which lead to either deficiency or functional abnormality of FIX. The available data (Giannelli et al., 1992) clearly indicate that haemophilia B is highly heterogeneous at the molecular level. Therefore, the methodology for elucidating the molecular pathology of FIX gene underlying haemophilia B must be able to screen the gene regions of likely functional significance in a simple and rapid way, thus providing the basis for the establishment of genotype/phenotype correlations. The development of polymerase chain reaction (PCR)-based methods of mutation detection facilitated the identification of genetic defects associated with haemophilia B.

In the present study, we report the application of single-strand conformation polymorphism (SSCP) analysis (Orita et al., 1989) to the screening of seven functionally important FIX gene regions (2.66 kb) in 9 unrelated haemophilia

B patients. In eight of the patients analysed, a sequence alteration was observed. Two of these are previously undescribed mutations (FIX_{Barcelos}, FIX_{Louisa}). We have also established the haplotypes linked to FIX gene in the haemophilia B patients, by using four intragenic polymorphisms. Molecular modeling studies were performed in order to predict the alterations introduced by the mutations in the structure of the FIX molecule.

MATERIALS AND METHODS

Subjects

Patients of different degrees of severity and relevant family members from 9 unrelated haemophilia B families were examined. Seven families were of Portuguese origin (including four severe, two moderate, and one mild haemophilia) and two of African origin (one severe and one mild hae-

Received November 17, 1992; accepted February 18, 1993.

*To whom reprint requests/correspondence should be addressed.

TABLE 1. Search for FIX Gene Mutations by PCR-SSCP Analysis

Region/ Exon	PCR product nt	Annealing temp. °C	Exon bp	Intron bp	5' UT ^a bp	3' UT ^a bp	Digestion enzyme	Fragment size
A	285	53	117	42	30	—	—	—
B/C	536	49	164/25	78/188/81	—	—	<i>FokI</i>	315/221
D	247	53	114	76/57	—	—	—	—
E	266	49	129	68/69	—	—	—	—
F	355	66	203	85/67	—	—	<i>AvaI</i>	239/116
G	276	59	115	111/50	—	—	—	—
H	688	60	623	77	—	65	<i>TaqI</i> <i>XhoII</i>	300/258/130 404/284

^a5' and 3' untranslated sequence.

mophilia). Factor IX:C was determined in an one-stage clotting assay by standard methods. Factor IX antigen (FIX:Ag) levels were determined by an ELISA assay using the kit from Diagnostica Stago (Parquet-Gernez et al., 1984).

DNA Isolation

Genomic DNA was extracted from peripheral blood leucocytes by a simple salting-out procedure (Miller et al., 1988).

In Vitro DNA Amplification

A total of 2.66 kb were amplified by PCR, including the entire coding sequence (1,383 bp), flanking intronic sequences (1,115 bp), a segment of the putative promoter, and the untranslated leader sequence: region A = -126 to 159, 285 bp; region B/C = 6,247 to 6,783, 536 bp; region D = 10,315 to 10,562, 247 bp; region E = 17,600 to 17,866, 266 bp; region F = 20,277 to 20,632, 355 bp; region G = 29,927 to 30,203, 276 bp; region H = 30,744 to 31,432, 688 bp (numbering system according to Yoshitake et al., 1985). These regions were amplified with 20-mer oligonucleotides as primers.

SSCP Analysis

Three hundred nanograms of genomic DNA were amplified in a reaction mixture containing 300 ng of each primer, 0.2 mM of each dNTP, 50 mM KCl, 10 mM Tris-HCl pH 7.2, 1.5 mM MgCl₂, 0.1% gelatin, 1U *Taq* DNA polymerase and 0.2 µl of [α -³²P]dCTP (3,000 Ci/mmol, 10 mCi/ml) in a volume of 100 µl. The annealing temperatures are indicated in Table 1. Whenever necessary, the amplification products were digested with the appropriate restriction endonuclease, prior to the SSCP analysis: fragment B/C with *FokI*, fragment F with *AvaI* and fragment H with *TaqI* or *XhoII* (see Table 1). A portion of the digested PCR products (2.5 to 5.0 µl) was mixed with 5.0 to 7.5 µl of a loading solution (80% formamide,

0.1% bromophenol-blue, 0.1% xylene cyanol, 1 mM EDTA, and 10mM NaOH) and heated at 90°C for 3 min. The samples were then applied to a 4.5% polyacrylamide gel (acrylamide/N',N'-methylene bis acrylamide, 39/1; 30 × 40 × 0.04 cm). Electrophoresis was performed in the cold room (4°C) at 40 W constant power. The gel was dried on a vacuum slab gel dryer, exposed overnight at -80°C, and autoradiographed.

Direct Genomic Sequencing

The Nusieve-agarose gel purified amplification products were resuspended in TE buffer. The PCR amplification primers were used as sequencing primers (Dodé et al., 1990).

Haplotype Analysis

The following FIX gene polymorphisms were analysed by PCR: *HinfI* intron 1 (Yoshitake et al., 1985), *TaqI* intron 4 (Camerino et al., 1984), *XmnI* intron 3 (Winship et al., 1984); Malmo allele = *MnlI* Exon F (McGraw et al., 1985). The amplifications were performed as previously described by Koerber et al. (1990). The *MnlI* polymorphism in exon F was analysed by SSCP.

Molecular Modeling of the FIX Mutations

The molecular model of the human factor IXa protease domain was built according to the structural homology it shows with other serine proteases, namely trypsin and elastase. Computational results were obtained using software programs from Biosym Technologies of San Diego. The molecule was minimised using the *Discover* programme. The mutations were created with the *Biopolymer* module of the *Insight II* and the model visualised through the graphics of *Insight II*. Secondary structure prediction was performed using the Chou-Fasman method (Chou and Fasman, 1978) included in the *Homology* module of *Insight II*.

RESULTS

We have applied SSCP analysis to the screening of functionally important FIX gene regions in 9 unrelated haemophilia B patients of Portuguese or African origin. Seven regions of the FIX gene were analysed to a total length of 2.66 kb (see Table 1). Eight of the patients displayed a mobility shift in one of the fragments analysed, indicating a sequence variation. Representative results of the SSCP experiments are shown in Fig. 1. Direct sequencing of the fragments with abnormal migration patterns led to the characterization of a sequence change in all positive cases (see Fig. 2). The results of the SSCP analysis, the sequence changes, and the haplotypes upon which the mutations occurred are summarized in Table 2.

No mobility shift was found in sporadic haemophilia B patient no. 141, with a FIX:C level of 25%. In the severely affected patient of African origin (no. 119) we identified a G to T transversion at position 6,596 in intron 2 (GTATGTTAA → GTATTTTAA). This base substitution is located 117 bp downstream the exon 2/intron 2 junction, which makes it more likely to represent a rare normal variant. The absence of a functionally significant mobility shift in these two patients might be explained by the slight insensitivity of the SSCP methodology. On PCR-SSCP analysis, we could detect the molecular basis of haemophilia B in 7 of the 9 analysed patients. This is in agreement with the sensitivity of this methodology reported by Orita et al. (1990). It can be further improved either by using different restriction fragments for the SSCP analysis or different electrophoretic conditions (Michaud et al., 1992).

The A to C transversion in codon 380 (FIX_{Barcelos}) creates a new Mva I restriction site. The presence of the mutation was thus confirmed in the patient's mother and affected brother. A different pattern of migration of single-stranded and double-stranded DNA fragments was observed in FIX_{Lousada} (see Fig. 1D, lane 7). The direct sequencing of this fragment revealed a 9 bp duplication at nucleotide 31,309 or 31,318. Six of the causative mutations presented here are located in the protease domain of FIX and one in the activation peptide region. Four of them occur at CpG sites, a hotspot for mutation in the FIX gene, two of which affecting the same amino acid (Arg 333).

The distribution of FIX mutations among haplotypes parallels the relative haplotype frequencies found in the Portuguese population (D. David, un-

published data). The detection of the Thr-148-Ala (Malmo) dimorphism by SSCP represents a further methodological improvement, since it allows the simultaneous examination of the Mnl I polymorphic site and surrounding DNA sequences in the same experiment (Fig. 1B).

DISCUSSION

The data presented in this study are the first FIX gene mutations identified in haemophilia B patients of Portuguese and African origin. Two of the seven presently identified mutations had never been previously reported.

The A to C transversion in codon 380 (ACT → CCT) for FIX gene identified in the patient with FIX_{Barcelos} is different from the previously reported mutations at this codon located just before the highly conserved region CR7 (Furie et al., 1982; Giannelli et al., 1992). The molecular model of FIXa indicated the presence of a hydrogen bond between the carbonyl oxygen of the Thr 380 and the amide nitrogen of Val 401. Secondary structure prediction suggests that substitution of proline for threonine at residue 380 enhances the probability of a turn in this region of the protein and may disallow the formation of the hydrogen bond to Val 401. This distortion in local conformation in the core of the molecule may adversely affect protein folding, leading to low antigen level (FIX:Ag = 1%).

Molecular modeling studies indicate that the 9 bp duplication (Ile-Tyr-Thr) in FIX_{Lousada} creates a new surface loop, which cannot easily be accommodated into the overall structure by energy minimisation.

Recent studies clearly indicate that the local sequence environment plays an important role in the generation of short deletions/insertions (Krawczak and Cooper, 1991; Cooper and Krawczak, 1991). We examined this region of the FIX gene looking for repetitive elements. A symmetrical heptamer in the immediate vicinity of the insertion and a direct repeat were found (see Fig. 3A). Two other deletion/insertion rearrangements have been reported in the same region (nucleotides 31,300–31,339). The sequence in this region was suggestive of a hypothetical model for the generation of the 9bp duplication based on the formation of an intermediate Moebius loop-like structure (Cooper and Krawczak, 1991) (see Fig. 3B). This loop could serve as an intermediate to potentiate the insertion. Breakage and partial resolution of one of the mismatched regions could make the loop a single/double-stranded structure, a hot spot

FIGURE 1. Representative results of SSCP analysis of several functionally important FIX gene regions. (A) Exon B/C: lane 1, patient 119 C→A intron 2. (B) Exon F: lane 1 *FIX*_{Porto} Arg-145-His, lanes 2–8, *Mnl* I allele +, lanes 9, 10, *Mnl* I

allele-. (C) Exon G: lane 1, *FIX*_{Louisa} Gly-207-Arg. (D) Exon H, *Taq* I digestion: lanes 1, 2, *FIX*_{Cascais} Arg-333-Stop, lanes 5, 6, *FIX*_{Barcelos} Thr-280-Pro, lane 7, *FIX*_{Louisa} Ins 9nt 31,300 or 31,309, lane 11, *FIX*_{Sesimbra} Arg-333-Gln.

for the DNA polymerase repair activity. Complete resolution of this structure would result in the 9-bp duplication. The presence of the inverted repeat ATACC /GGTAT could suggest an alternative model, through the formation of an intra-strand

loop. Breakage of this loop after the ATACC motif followed by slipped mispairing and resynthesis of the 9bp sequence could also explain the duplication.

The analysis of the distribution of short dele-

FIGURE 2. Partial nucleotide sequences of the fragments with abnormal migration patterns, corresponding to the two novel mutations. The arrowheads indicate the site of the mutation. (A) FIX_{Buenos} Thr-380-Pro (sense strand). (B) FIX_{London} Ins 9 bp (antisense strand).

tions/insertions in the FIX gene (Giannelli et al., 1992) revealed two additional clusters of this type of rearrangement, at nucleotides 6,398–6,402 and 31,141–31,116, respectively. The concentration of repetitive DNA sequences and the presence of symmetrical elements in those regions, is consistent with the interpretation that mutational lesions are strongly influenced by the DNA sequence environment (Krawczak and Cooper, 1991).

As found for the patient with FIX_{Penafiel} (Arg-248-Gln), other moderate to mild haemophilia B

phenotypes have been reported which result from the replacement of the same Arg-248 residue (Giannelli et al., 1992). Arg-248 is located on the surface of the molecule, close to the putative Ca^{2+} -binding surface loop, a region involved in the binding of the protein cofactors essential for FIX action (Bajaj et al., 1992). Loss of a positive charge from the surface of the molecule is not consistent with protein stability and could explain the low antigen level observed in this patient (FIX:Ag = 3%).

Region	Patient	Nucleotide position	Nucleotide change	Codon	Codon change	FIX:C (%)	FIX:Ag (%)	HinfI	TaqI	XmnI	MnII
Exon F	Porto	20,414	G→A	145	Arg→His	4	64	-	-	-	-
Exon G	Luanda ^a	30,072	G→A	207	Gly→Arg	1	ND	-	+	+	+
Exon H	Penafiel	30,864	G→A	248	Arg→Gln	3	3	-	+	+	+
	Sesimbra	31,119	G→A	333	Arg→Gln	<1	ND	-	-	-	-
	Cascais	31,118	C→T	333	Arg→Stop	1	1	+	-	-	-
	Barcelos	31,259	A→C	380	Thr→Pro	<1	1	-	-	-	-
	Lousada	31,309 or 31,318	Ins 9nt	399	In frame ^c	1.5	2	-	-	-	-
Intron 2	119 ^b	6,596 IVS2, -117	G→T	—	—	1	1	-	-	-	-
	141	—	—	—	—	25	ND	+	-	-	-

^cInsertion of Ile-Tyr-Thr. ND, not determined.

*FIXLousada

FIGURE 3. (A) Nucleotide sequence of the hot spot for short deletion/insertions in the factor IX gene between nucleotides 31,300 and 31,339. The sequence in bold indicates a 7-bp symmetrical element and the underlined sequences indicate other repetitive elements such as direct repeats. (B) Hypothetical model for the generation of the 9 bp duplication (FIX_{Lambton}) based on the formation of a Moebius loop-like intermediate secondary structure.

A severe form of haemophilia B resulting from the replacement of Gly-207 similar to FIX_{Laurada} has been reported (Giannelli et al., 1992). The Gly-207 is situated in the core of the molecule in the first disulphide loop of the catalytic domain (Sarkar et al., 1990). The mutation to a bulky arginine residue causes very severe nonbonded in-

teractions, probably preventing the correct folding of this molecule.

Several recurrent nonsense and missense mutations were reported at codons 333 and 145 as in FIX_{Scimbra} Arg-333-Stop, FIX_{Cascais} Arg-333-Gln and FIX_{Porto} Arg-145-His, respectively (Giannelli et al., 1992). Arg-333 lies on the surface of the

molecule in the FIX-specific polypeptide region (Sarkar et al., 1990). The analogous mutation in factor VIIa abolishes binding to tissue factor, the co-factor of factor VIIa (O'Brien et al., 1991). The Arg-145 is one of the two sites at which factor XIa activates single chain FIX to two-chain activated form.

ACKNOWLEDGMENTS

This research was partially supported by research grant 234.90/SAU from Junta Nacional de Investigação Científica e Tecnológica (JNICT). H.A.V. Rosa was supported by a research scholarship from JNICT. We are grateful to the clinicians, especially to Dr Sara Morais for referring to us their patients, and to Dr Maria Irene Pereira and Dr Graça de Deus for performing the FIX:Ag levels on the plasma samples from these patients. The authors thank Prof. David Cooper for his critical reading of the manuscript and helpful comments.

REFERENCES

- Bajaj SP, Sahharwal AK, Gorka J, Birktoft JJ (1992) Antibody-probed conformational transition in the protease domain of human factor IX upon calcium binding and zymogen activation: Putative high affinity Ca^{2+} -binding site in the protease domain. *Proc Natl Acad Sci USA* 89:152-191.
- Camerino G, Grieschik KH, Jaye M, De La Salle H, Tolstoshev P, Lecocq JP, Heilig R, Mandel JL (1984) Regional localization on the human X chromosome and polymorphism of the coagulation factor IX gene (hemophilia B locus). *Proc Natl Acad Sci USA* 81:1771-1773.
- Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol* 47:45-147.
- Cooper DN, Krawczak M (1991) Mechanism of insertional mutagenesis in human genes causing genetic disease. *Hum Genet* 87:409-415.
- Dodé C, Rochette J, Krishnamoorthy R (1990) Locus assignment of human α globin mutations by selective amplification and direct sequencing. *Br J Haematol* 76:275-281.
- Furie B, Bing DH, Feldmann RJ, Robinson DJ, Burnier HP, Furie BC (1982) Computer-generated models of blood coagulation factor Xa, factor IXa, and thrombin based upon structural homology with other serine proteases. *J Biol Chem* 267:3875-3882.
- Giannelli F, Green PM, High KA, Sommer S, Lillicrag DP, Ludwig M, Olek K, Reitsma PH, Gonssens M, Yoshioka A, Brownlee GG (1992) Haemophilia B: database of point mutations and short additions and deletions—third edition, 1992. *Nucl Acids Res* 20:2027-2063.
- Krawczak M, Cooper DN (1991) Gene deletions causing human genetic disease: Mechanisms of mutagenesis and role of the local DNA sequence environment. *Hum Genet* 86:425-441.
- Koeberl DD, Bottema CDK, Sarkar G, Ketterling RP, Chen S-H, Sommer SS (1990) Recurrent nonsense mutations at arginine residues cause severe hemophilia B in unrelated hemophiliacs. *Hum Genet* 84:387-390.
- McGraw RA, Davis LM, Noyes CM, Lundblad RL, Roberts HR, Graham JB, Stafford DW (1985) Evidence for a prevalent dimorphism in the activation peptide of human coagulation factor IX. *Proc Natl Acad Sci USA* 82:2847-2851.
- Michaud J, Brody LC, Steel G, Fontaine G, Martin LS, Valle D, Mitchell G (1992) Strand-separating conformational polymorphism analysis: Efficacy of detection of point mutations in the human ornithine γ -aminotransferase gene. *Genomics* 13:389-394.
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting-out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 16:1215.
- O'Brien DP, Gale KM, Anderson JS, McVey JH, Miller GJ, Meade TW, Tuddenham EGD (1991) Purification and characterization of factor VII 304-Gln: A variant molecule with reduced activity isolated from a clinically unaffected male. *Blood* 78:132-140.
- Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5:874-879.
- Orita M, Sekiya T, Hayashi K (1990) DNA sequence polymorphisms in A16 repeats. *Genomics* 8:271-278.
- Parquer-Gomez A, Mazurier C, Amihai J, Martinoli JL (1984) Assay of factor IX antigen using an enzyme immuno assay. *Thromb Res* 35:703-712.
- Sarkar G, Koeberl DD, Sommer SS (1990) Direct sequencing of the activation peptide and the catalytic domain of the factor IX gene in six species. *Genomics* 6:133-143.
- Winship RP, Anson DS, Rizza CR, Brownlee GG (1984) Carrier detection in haemophilia B using two further intragenic restriction fragment length polymorphisms. *Nucl Acid Res* 12:8861-8872.
- Yoshitake S, Schach BG, Foster DC, Davie EW, Kurachi K (1985) Nucleotide sequence of the gene for human factor IX (anti-hemophilic factor B). *Biochemistry* 24:3736-3750.