Severe haemophilia B due to a 6 kb factor IX gene deletion including exon 4: Non-homologous recombination associated with a shortened transcript from whole blood

Ting-Chang Hsu, Shelley M. Nakaya, Arthur R. Thompson
Puget Sound Blood Center and the University of Washington, Seattle, Washington, USA

Summary
In genotyping a severe haemophilia B subject, exons 1–3 and 5–8 were normal. Exon 4 did not amplify, suggesting a partial gene deletion. Previously, a French family with an exon 4 deletion had severe haemophilia B with a circulating, dysfunctional factor IX protein missing its first growth factor-like domain; breakpoints were not analyzed. Using a 5’ primer for exon 3 and a 3’ primer for exon 5 fragments, the subject’s factor IX gene amplified a 5 kb fragment whereas 11 kb was predicted, indicating a 6 kb deletion. Restriction endonucleases localized the 3’ intron 4 deletion breakpoint to 1.2 kb 5’ to exon 5. Sequencing through the breakpoints revealed a 5,969 bp deletion that included exon 4 and was accompanied by a 13 bp duplication inserted near the 3’ breakpoint site. Haemophilia was familial; on testing, his mother was confirmed as a heterozygous carrier, whereas his sister was homozygous for the normal, larger fragments. As exons 4 and 5 of the factor IX gene are in frame, this deletion should produce a shortened transcript, missing 114 bp (38 codons from the first growth factor-like domain). Reverse transcription of mRNA prepared from whole blood and PCR identified the shorter cDNA fragment. Western blotting demonstrated a smaller factor IX protein.

Keywords
Deletion breakpoints, ectopic mRNA, exon 4, factor IX gene, haemophilia B

Introduction
Large deletions of the factor IX gene are found in about 5% of families with severe haemophilia B. About 100 have been reported and half of these are partial (1). Only one has been restricted to the 4th exon (2). In addition, there have been four deletions that were complex with evidence of a large insertion or inversion accompanying the deletion event (1). To date, only seven of the partial F9 gene deletions have had breakpoints defined despite the complete sequence of the factor IX gene being available for over 20 years (3). One of these had homologous recombination between 14 bp repeat sequences in introns 4 and 7, deleting exons 5 and 6 (4). This deletion is notable as an amino-terminal protein fragment was expressed and recovered in the affected family member’s urine (5). The other six are a 9.3 kb deletion including exons 4 and 5 (6), a 1.1 kb deletion including exon 5 (7), a 6.5 kb deletion including exon 6 (8), a 1.4 kb deletion including exon 7 (1), an 8.8 kb deletion from intron 6 to the end of the 3’ untranslated sequence (7) and a 22.4 kb deletion including exons 7 and 8 (9). The latter case was the only one accompanied by a short duplication that was 16 bp.

Plasmas from three affected members of the French family with an isolated exon 4 deletion had a severe phenotype with no detectable factor IX clotting activity, although there were detectable levels of factor IX antigen (2). This was confirmed in a polyclonal factor IX ELISA; furthermore, immunoassays coating wells with a variety of monoclonal anti-human factor IX preparations confirmed that epitopes within the Gla domain, the activation peptide and the catalytic domain were present (10). However, a monoclonal antibody to an epitope within the first growth factor-like domain reacted with normal but not the haemophilic factors IXs or IXa light chains (10).
Detection of carriers of large, partial deletions is problematic as standard PCR reactions are not sufficiently quantitative to distinguish a single from two template copies in a woman’s genome. Patterns on Southern blots have been successful as well as the detection of an abnormal protein fragment in urine (5), but direct analysis by PCR has been largely limited to the few families in which the breakpoints have been defined. Primers used on genomic DNA readily distinguished a haemophilic 5 kb fragment in the currently reported family from a normal 11 kb fragment. The breakpoint was then identified by sequencing. Reverse PCR of lymphocyte-derived ectopic transcription of factor IX mRNA has been problematic as no product could be obtained for sequences provided by exons 1–6 (11) or 3–6 (12). More recently, however, a product spanning sequence provided by the 3rd exon was obtained (13). In the current study, a cDNA was generated from a whole blood sample using primers from sequences within the 2nd and 5th exons that allowed demonstration of the patient’s fragment that was missing 114 bp.

**Methods**

**Subject and family**

The propositus is an 18-year-old male with severe haemophilia B diagnosed in infancy (<1% factor IX clotting activity performed elsewhere). He has been on prophylactic infusions of factor IX concentrate since early childhood and is a competitive college athlete. A maternal uncle has severe haemophilia B. Samples were obtained from his mother to confirm her carrier state and from his sister to determine her carrier status. Samples were drawn after informed consent according to a protocol approved by the University of Washington’s Institutional Review Board.

**Genomic DNA**

DNA was extracted from peripheral blood leukocytes as previously described, as were amplification primers and conditions (14). For the numbering of base pairs, the GenBank entry NC_000023.9, March, 2006 for the X chromosome of the human genome was used; for comparison, the sequence of Yoshitake et al. (3) is provided where indicated. A new intron 4 antisense primer, GAGCTAAGTGTAAAGACCT, representing bases 138455975–138455956 was used to localize the deletion breakpoint. It was paired with an intron 2 sense primer, agataggaaatcaataa (bases 138447120–39), previously synthesized as a 5′ sense primer for exon 3. Restriction endonucleases were purchased from New England Biolabs (Ipswich, MA, USA) and digests were performed according to the package insert followed by gel electrophoresis in agarose (14).

To demonstrate the deletion, genomic DNA was PCR amplified using an Expanded Long Range Template System (Roche, Nutley, NJ, USA). Amplification conditions were 94°C for 2 minutes (min), then 10 cycles of 94°C 10 seconds (sec), 50°C 30 sec, and 68°C 10 min, followed by 20 cycles where 10 sec was added to the final 68°C incubation of each cycle, and ending with an additional 7 min at 68°C. DNA sequencing was performed on an Applied Biosystems model 3100 capillary sequencer (Foster City, CA, USA) using amplification primers. The breakpoint sequence was confirmed by a separate sequence of a second amplified fragment from the patient’s genomic DNA.

**RNA preparation from whole blood**

Total RNA from three normal subjects and the haemophilic subject with a deletion were prepared using PureLink™ Total RNA Blood Purification Kit (Invitrogen, Carlsbad, CA, USA) to isolate total RNA from whole blood. During the procedure, RNase AWAY Reagent (Molecular BioProducts, San Diego, CA, USA) was used to remove RNase contamination from surfaces. A 500 µl aliquot of an EDTA anticoagulated whole blood sample was used for each preparation according to the Invitrogen Instruction Manual. Total RNA was eluted from the kit’s Spin Cartridge by 100 µl of RNase free water. The RNA concentration was measured by a UV spectrometer with Absorbance at 260 nm.

**Synthesis of first strand cDNA**

First strand of partial cDNA of factor IX was synthesized using the StrataScript First-Strand Synthesis System (Stratagene, La Jolla, CA, USA). A specific primer was used based on factor IX cDNA sequence 3′ to the exon 4 deletion (CACTGCTGGTTACAGG; antisense, beginning with the first two bases from codon 128 in exon 6 and the last 15 bp coded for by exon 5). About 200 µg total RNA was used in the reverse transcriptase reaction according to the Stratagene kit manual. Immediately after the reaction, PCR was performed using a 5′ primer based on sequence from exon 2 of the factor IX gene (GGTAAATTGGAAGAGTTTGTT; sense, representing codons 4 through 10 in the mature protein) and the 3′ primer from the reverse transcriptase reaction (above). The PCR conditions were: 94°C 3 min, then 35 cycles of 94°C for 1 min, 58°C for 1 min and 72°C for 2 min followed by a final incubation of 7 min at 72°C.

**Figure 1: Amplified fragments from intron 2 into intron 4; family study.** The 2nd through the 5th lanes from the left are fragments from a normal subject, the haemophilic subject, his mother and sister, respectively. Genomic DNA was amplified with a 5′ primer from intron 2 and the “new” internal 3′ intron 4 primer. The subject’s fragment is 6 kb smaller than the normal, whereas his mother is heterozygous as predicted for a carrier. His sister is homozygous for the normal factor IX gene fragments and is thus not a carrier. The left lane contains standard sized fragments; black arrows with numbers refer fragment sizes in kb. Samples were electrophoresed on a 1% agarose gel.
Western blotting

Plasmasamples from the propositus and a normal control subject were diluted 1:10 in 20mM Tris-HCl pH 7.0, 150mM NaCl (TBS). Samples were electrophoresed on 7.5% polyacrylamide gels with SDS and blotted to nitrocellulose. Blots were blocked with 4% BSA containing TBS with 0.2% Triton X-100 (TBST) and then incubated with peroxidase-conjugated goat anti-human factor IX (Affinity Biologics, Ancaster, ON, Canada) diluted 1:1,000 ratio in TBST from a 1mg/ml stock solution. Blots were washed with TBST three times, then stained with ECL Western Blotting Detection Reagents (Amersham Biosciences, Piscataway, NJ, USA) and exposed on an X-ray film.

Results

In initial studies on the haemophilic subject’s genomic DNA, all factor IX exon fragments were amplified and normal on size and sequence except for exon 4 where repeated attempts to amplify showed no detectable product. This suggested that exon 4 was involved in a partial gene deletion. Using long-range PCR with a 5’ exon 3 sense primer and a 3’ exon 5 primer (14), a 5 kb fragment was amplified from the haemophilic subject’s genomic DNA, whereas no product was observed from the normal subject (not shown). As an 11 kb fragment is predicted from the genomic sequence, this suggested a 6 kb deletion including exon 4.

Using the amplification primers for DNA sequencing of the 5 kb haemophilic fragment, only normal sequence was demonstrated. Included were base 138449767, 756 bp 3’ to the end of exon 3 and base 138457538, 640 bp 5’ to the beginning of exon 5. The latter, in intron 4, was determined from the antisense primer for exon 5. This indicated that the 3’ breakpoint of a deletion occurred within the central portion of the 7.2 kb intron 4. To further localize the 3’ breakpoint, restriction digest patterns of the 5 kb haemophilic fragment determined from the genomic sequence (3) are 9287–15249 where intron sequences are slightly different from subsequent, automated results. Note that the proband is missing 138449797 through 138455765 or 5,969 bases. Additionally, there is a 13 bp duplication (boxed) that begins 6 bp 3’ to the 3’ breakpoint in intron 4 of the haemophilic gene. Along with the last base, an “a”, prior to the 5’ breakpoint in intron 3, the duplication occurs between two 7 bp repeat sequences. The diagram beneath the sequence data indicates the relative sizes of the eight F9 exons (e) and 3’ untranslated sequence and, on a different scale, the relative sizes of the seven introns, as adapted from Bajaj and Thompson (15). Arrows above represent the breakpoint sites in the 3.7 kb intron 3 and 7.2 kb intron 4, that encompass the subject’s 6 kb deletion.

Figure 2: Sequence through the deletion breakpoint. Sequence is shown in the sense orientation as determined from an antisense primer in intron 4. The ends of the deleted sequence are in italics and underlined to show the breakpoints; bases are numbered according to the GenBank sequence NC_000023.9. The corresponding sequence numbers for the deletion in the 1985 Yoshitake sequence (3) are 9287–15249 where intron sequences are slightly different from subsequent, automated results. Note that the proband is missing 138449797 through 138455765 or 5,969 bases. Additionally, there is a 13 bp duplication (boxed) that begins 6 bp 3’ to the 3’ breakpoint in intron 4 of the haemophilic gene. Along with the last base, an “a”, prior to the 5’ breakpoint in intron 3, the duplication occurs between two 7 bp repeat sequences. The diagram beneath the sequence data indicates the relative sizes of the eight F9 exons (e) and 3’ untranslated sequence and, on a different scale, the relative sizes of the seven introns, as adapted from Bajaj and Thompson (15). Arrows above represent the breakpoint sites in the 3.7 kb intron 3 and 7.2 kb intron 4, that encompass the subject’s 6 kb deletion.

Hsu et al. Exon 4 deletions in severe haemophilia B

Figure 2: Sequence through the deletion breakpoint. Sequence is shown in the sense orientation as determined from an antisense primer in intron 4. The ends of the deleted sequence are in italics and underlined to show the breakpoints; bases are numbered according to the GenBank sequence NC_000023.9. The corresponding sequence numbers for the deletion in the 1985 Yoshitake sequence (3) are 9287–15249 where intron sequences are slightly different from subsequent, automated results. Note that the proband is missing 138449797 through 138455765 or 5,969 bases. Additionally, there is a 13 bp duplication (boxed) that begins 6 bp 3’ to the 3’ breakpoint in intron 4 of the haemophilic gene. Along with the last base, an “a”, prior to the 5’ breakpoint in intron 3, the duplication occurs between two 7 bp repeat sequences. The diagram beneath the sequence data indicates the relative sizes of the eight F9 exons (e) and 3’ untranslated sequence and, on a different scale, the relative sizes of the seven introns, as adapted from Bajaj and Thompson (15). Arrows above represent the breakpoint sites in the 3.7 kb intron 3 and 7.2 kb intron 4, that encompass the subject’s 6 kb deletion.

Of three new 3’ (antisense) primers from intron 4, the one indicated in methods is just 5’ to the PsII site and was used to amplify fragments shown in Figure 1, paired with an intron 2, 5’ sense primer for exon 3. When the new intron 4 antisense primer was used as a sequence primer, the intron 3 and intron 4 deletion breakpoints were demonstrated as shown in Figure 2. The deletion is 5,969 bp including bases 138449797–138455765 with a 13 bp insertion (Fig. 2). The insertion is a duplication of AGGGGTACATGTG that occurs 6 bp 3’ to the 3’ deletion breakpoint in intron 4 between two identical 7 bp sequences (AAGTTT) in the haemophilic factor IX gene. Note that the first base of the 5’ identical sequence is “A” just prior to the 5’ deletion breakpoint in intron 3.

Using the new intron 4 antisense primer with the 5’ primer from exon 3 and the long-range PCR conditions, a normal fragment of 8,856 bp is predicted, whereas the haemophilic fragment is predicted as 2,887 bp. With these amplifications, the subject’s
mother was confirmed as a carrier, whereas his sister was homozygous for the normal sized band and thus is not a carrier (Fig. 1).

The products of RT-PCR reactions from the normal and haemophilic subjects were analyzed. The predicted sizes of amplified cDNA are 375 bp for the normal subject and 261 bp from the haemophilic subject. The observed difference (Fig. 3) corresponds to the 114 bp (38 codons) of exon 4 of the factor IX cDNA. Sequencing confirmed the predicted result.

On Western blots, the plasma from the propositus showed both a normal sized factor IX species from his prophylactic therapy and a smaller species corresponding to his haemophilic factor IX, predicted to be ~4,000 KD lower molecular weight (Fig. 4).

Discussion

The factor IX gene (F9) is 34 kb and composed of 8 exons, the 4th of which codes for the first of two growth factor-like domains (3, 15). As the 4th exon begins and ends in frame, one would predict that a partial gene deletion that only includes that exon would code for an mRNA that has 114 bp shorter than normal and that a protein would be expressed that has that domain deleted. In the crystal structures that have been solved for factor IXa, the first growth factor-like region is predicted to extend the catalytic domain at a proper height above a lipid membrane to support intrinsic X-ase (15). A shortened factor IX would be unable to participate in clotting on biologically active membranes, even though it can undergo activation cleavages.

In a French family with a partial factor IX gene deletion including exon 4, occurrence was de novo and Southern blotting estimated a 2.8 kb extent (2). Thus, the deletion was an independent event and included at least one different breakpoint than that found for the current subject even though each had an isolated deletion of exon 4. Our previous characterization of their patients’ plasmas confirmed that they had detectable, circulating factor IX protein. In addition, we found that their haemophilic factor IX could undergo activation cleavages by factor XIa, had a smaller than normal light chain on Western blots, and failed to react with a murine monoclonal antibody known to bind to an epitope within factor IX’s first growth factor-like domain (10); additional details of these results are available on request. Similar results have been obtained in expression of a factor IX cDNA in which the same sequence has been deleted (16). A similar study on the current subject’s plasma factor IX was unsuccessful due to his frequent, high-dose, prophylactic factor IX infusions. Infused factor IX would not affect his F9 mRNA, however. Accordingly, an appropriately shortened transcript was identified from ectopic mRNA in his leukocytes, confirming exon 4 “skipping” at the mRNA level. This would account for the smaller protein as observed in both the French family’s affected members (10) as well as that generated in vitro (16). This represents the first demonstration of an ectopic transcript including factor IX’s 4th exon as previous attempts from leukocytes had only identified more 3’ sequences (11, 12). The current results, along with a more recent demonstration of a transcript including exon 3 (13), suggest that previous negative results were likely due to technical limitations.

Large partial or even complete gene deletions have been described in haemophilia B (1). Of the few where the breakpoint has been analyzed, some have been between homologous sequences. In the current family, the deletion breakpoints were in non-homologous sequences. The genotype is more complex than a large deletion, however, as a small, 13 bp insertion occurred as a duplication between identical 7 bp segments. The fact that the first base of the 5’ segment is from intron 3 suggests that the duplication arose after or as a result of the large deletion. In the one other case with a small insertion, 16 bp that were not duplicated joined the two breakpoints (9). Although that recombination was
between non-homologous sequences, there was a 9 bp repeat in the opposite orientation in the preserved sequences near each breakpoint.

In a database for breakpoints of large gene deletions and translocations in a variety of species, non-homologous recombinations were more frequent in the latter (17). Furthermore, non-homologous deletions were more often flanked by AT-rich and translocations with GC-rich alternating purine-pyrimidine sequences within up to 30 bp of the breakpoint. In the current deletion, neither purine-pyrimidine tract is present at either breakpoint. Among deletions, 80 of 93 occurred within repetitive elements (17). Data are based primarily on non-human gene rearrangements, and there have been insufficient analyses among deletions in patients with haemophilia A or B to determine the degree to which these monogenic disease-causing events are similar to those across the wide variety of microorganism, plant and non-human species. Characterization of breakpoints in additional patients as in the current subject are needed to determine if there are common features to suggest mechanisms involved in the etiologies of large deletion genotypes leading to severe haemophilia.

Carrier detection by direct genomic analysis is clearly preferable to linkage as fewer samples are required and false paternity is not an issue. Southern blotting has been effective in demonstrating heterozygosity of partial gene deletions as TaqI digestion of genomic DNA hybridized to a labeled factor IX cDNA provides specific bands for the different exons. Alternatively, detection of a truncated protein can sometimes be used (4, 5) or RT-PCR, although the latter have failed to demonstrate amino-terminal ectopic factor IX transcripts in RNA from leukocytes (11, 12). Attempts to quantify amplification of different exons can also distinguish carriers (18, 19). Each of these procedures can be time consuming and difficult to standardize and perform. As shown in the current family, once the extent of involvement of specific introns is known, simple, direct genomic analysis of an amplified fragment can demonstrate or exclude heterozygosity for the deletion and is thus a practical application of studies to define a partial gene deletion.

Western blots confirmed the presence of a smaller factor IX species in the propositus. The haemophilic species appeared less concentrated than either the native factor IX in normal plasma or from his prophylactic infusion in his own plasma. Although not quantitated, his haemophilic factor IX appears in a comparable amount to the 30% antigenic level in the French family (2). In the previous family with an exon 4 deletion, it was also shown that, following factor XIa activation cleavages, there was a smaller amino-terminal light chain (10). Thus, in both families, an isolated exon 4 deletion is associated with a circulating, non-functional, somewhat lower molecular weight haemophilic factor IX protein.

Acknowledgements

The subject and his family are thanked for their generous participation. This study was supported in part by NIH R01HL071093–01A2, and by a Judith Graham Pool Fellowship from the National Hemophilia Foundation for Dr. Hsu.

References