Molecular characterization of hemophilia B in North Indian families: identification of novel and recurrent molecular events in the factor IX gene

[haematologica]
2004;89:1498-1503

ANUBHA MAHAJAN
SREENIVAS CHAVALI
MADHULIKA KABRA
MADHUMITA ROY CHOWDHURY
DWAIPAYAN BHARADWAJ

A B S T R A C T

Background and Objectives. Hemophilia B is an X-linked recessive, bleeding disorder caused by mutations in the factor IX gene. A wide range of mutations, showing large molecular heterogeneity, has been described in hemophilia B patients. Our study was aimed at characterizing mutations in the factor IX gene in a cohort of North Indian hemophilia B patients.

Design and Methods. Polymerase chain reaction (PCR) amplification and direct sequencing of all regions of likely functional significance- the coding regions, promoter, the 5' UTR, the splice junctions and parts of the 3' UTR of the factor IX gene was done in 18 families carrying a severe form of hemophilia B.

Results. We identified 10 point mutations (including 2 novel ones); one novel deletion and one donor splice site mutation. Recurrence of a nonsense and a missense mutation was observed. The mutation in 3 families could not be characterized. None of the 14 polymorphic positions reported in the *Haemophilia B Mutation database* in the regions sequenced were polymorphic; herein we report four novel synonymous single base mismatches. One mutation reported to be causative in the database was found to be more likely a non-causal polymorphism.

Interpretation and Conclusions. Our data confirm the remarkable heterogeneity of the mutational spectrum in hemophilia B among affected families. This is the first mutation report on the disease in the Indo-Aryan population from the Indian subcontinent. Identification of a causative mutation leads to more precise carrier detection than does conventional polymorphism-based linkage analysis. This can effectively be used to establish genotype/ phenotype relationships.

Keywords: Factor IX, hemophilia B, molecular events, North Indian.

Functional Genomics Unit, Institute of Genomics and Integrative Biology, CSIR, Delhi, India (AM, SC, DB); Genetics Unit, Department of Pediatrics, All India Institute of Medical Sciences, New Delhi, India (MK, MRC).

Correspondence:
Dr. Dwaipayan Bharadwaj, Functional Genomics Unit, Institute of Genomics and Integrative Biology (CSIR), Mall Road, Delhi- 110 007, India.
E-mail: db@igib.res.in

Hemophilia B (MIM# 306900) is an X-linked, recessive bleeding disorder caused by mutations in the factor IX gene (*F9*; GenBank accession number: K02402.1). To date, a wide range of mutations, showing large molecular heterogeneity, has been described in hemophilia B patients, as can be seen from the Haemophilia B Mutation database.[1] Mutation analyses of *F9* of hemophilia B patients from different cohorts[2-4] emphasize this heterogeneity. The level of factor IX in hemophilia B varies among affected individuals and the disease is classified as severe, moderately severe or mild based on the normal factor IX activity assayed in the patients' plasma (<1, 2-5%, 6-30%, respectively). Factor IX is a trypsin-like serine protease functioning in the intrinsic pathway of blood coagulation. It is activated by either factor VIIa/ tissue factor complex or factor XIa in a calcium-dependent reaction. Once activated, factor IXa with its co-factor, factor VIIIa, in the presence of calcium and phospholipid forms the tenase complex and activates factor X. *F9* is located at Xq27.1 and consists of 8 exons coding for 6 functional domains, namely the prepropeptide, the Gla domain, epidermal growth factor (EGF)-1 and EGF-2 domains, activation domain, and the catalytic domain.[5]

Factor IX circulates as a 415 amino acid single chain zymogen with a molecular mass of 55kDa and is present in normal plasma at a concentration of approximately 5 μg/mL. Factor IX is synthesized in hepatocytes as a precursor containing an amino terminal signal sequence and a propeptide. The signal sequence and propeptide are cleaved off in separate reactions prior to secretion and failure to remove these sequences results in a non-functional pro-

Table 1. Primers for mutation analysis of factor IX gene.

Functional region	*Sequence of the primer*	*Product size (bp)*	*Annealing temperature in PCR (°C)*
Promoter and Exon 1	5' GGGAGATGGACATTATTT 3' 5' GTGAAGAAGACAGCATCAGAT 3'	275	58
Exon 2 & Exon 3	5' AAAACAAAGACTTTCTTAAGAGAT 3' 5' CATATGTTTCATATATTAGCTAGAG 3'	506	55
Exon 4	5' ATCCCAATGAGTATCTACAGGG 3' 5' GTTTCAACTTGTTTCAGAGGG 3'	254	62
Exon 5	5' CCATGTACTTTTTAGAAATGC 3' 5' GCTGAAGTTTCAGATACAGA 3'	242	55
Exon 6	5' GCCAATGAGAAATATCAGG 3' 5' TCTTGCCAGCTGAGCTCCAG 3'	296	60
Exon 7	5' GCCTATTCCTGTAACCAGCACAC 3' 5' GAGCTAGTGGTGCTGCAGAT 3'	242	65
Exon 8	5' TTAGGTCAGTGGTCCCAAGTAG 3' 5' TCGGTCAACAAGTGGAACTCTAAGG 3'	385	65
Exon 8	5' CTTCCTCAAATTTGGATCTGGC 3' 5' AGTGATTAGTTAGTGAGAGGCCCTG 3'	425	65
Poly A	5' GAACCGTTCGTTTGCAATCTACAG3' 5' GCCACTGGGCCCAGCCAAGAA 3'	584	59

tein.[6] The signal sequence directs the protein to the endoplasmic reticulum. Propeptide contains elements which are important for recognition of factor IX by vitamin K-dependent γ-glutamyl carboxylase, which is associated with the inner surface of endoplasmic reticulum.[7] The 38 residues amino terminal of the mature protein comprise a functional region known as the Gla domain, so named because its first 12 glutamate residues undergo a posttranslational vitamin K-dependent carboxylation at the γ-carbon. The Gla domain functions in calcium-dependent lipid binding and in binding to endothelial cells.[8] EGF-1 and EGF-2 act as calcium binding sites involved in platelet and activated factor VIII binding to factor IXa to form tenase complex. Cleavage of the activation peptide (residues 145-181) leads to activation of factor IX. The catalytic domain harbors the catalytic triad, His221, Asp269, Ser365, which is highly conserved in all serine proteases. Investigations have shown that hemophilia B can be caused by missense mutations resulting in amino acid substitutions; nonsense mutations leading to truncations; insertions or deletions causing frameshifts and mutations in the splice sites as well. Identification of the causative mutation provides many advantages compared to conventional polymorphism-based linkage analysis, leading to precise carrier and prenatal diagnosis. Herein, we report the molecular characterization of 18 unrelated patients of North Indian origin with severe hemophilia B.

Design and Methods

Patients of North Indian origin, affected by hemophilia B, were recruited from the All India Institute of Medical Sciences, New Delhi. Informed consent was obtained from each of the participants of this study. The disease status of the patients was established by clinical investigations which included the one stage-clotting assay that determines the severity of the disease. Eghteen families (HBFI-18) with severe hemophilia B were analyzed. For each patient at least two first or second degree relatives were also analyzed. In every case the mother of the patient was analyzed to establish the presence of the mutation in the heterozygous state.

Ten milliliters of blood were collected from each subject into tubes containing EDTA and leukocyte DNA was isolated as described elsewhere.[9] All regions of likely functional significance, including all the coding regions, promoter, the 5' UTR, the splice junctions and parts of the 3' UTR of the factor IX gene, were amplified by PCR using the primer sequences given in Table 1. Sequencing was done using an ABI PRISM 377 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The amplified products were purified from agarose gels using a gel extraction kit (Biological Industries, Kibbutz Beit Haemek, Israel). The purified products were then sequenced using both forward and reverse primers for

each amplified product, to ensure the presence of each mutation.

Results

All the regions of likely functional significance of the factor IX gene were sequenced in 18 hemophilia B families of North Indian origin. Each patient had only one amino acid altering sequence mismatch that was considered a candidate for a disease-causing mutation. These putative mutations were neither found as second site changes on analysis of any other hemophiliacs nor as neutral polymorphisms in healthy individuals of the same ethnicity. Sequencing of a second independent PCR product from each propositus and the carrier females belonging to the same family was always performed to validate mutations.

Screening for molecular events in the factor IX gene in these 18 families led to the characterization of ten point mutations, of which two were novel; one novel deletion of ten nucleotides and one donor splice site mutation (Table 2). Arg29Ter was observed in three families and Asn92Asp in two families. Mutations in the regions of the factor IX gene that we sequenced could not be characterized in 3 families. Four novel single base mismatches were detected in some subjects we analyzed (Table 3).

Discussion

A novel A→C transversion in exon 7 at nucleotide position 30105 (numbering as in Yoshitake S *et al.*)[10] was observed in HBF 2. This transversion leads to the substitution of Thr (a polar residue), by Pro (an imino acid) at position 218 in the catalytic domain. Thr218 is conserved in all species of factor IX and is part of a conserved region harboring His221 of the active center that is delimited by Cystine 206-222. Replacement by Pro might introduce a kink, thus affecting the orientation of His221 in the catalytic triad. Hence, this missense mutation causes the disease.

Another novel point mutation observed was a T→A transversion at nucleotide position 30829 in exon 8 of the factor IX gene (HBF 9). As a result of this mutation His236 of the catalytic domain is substituted by Gln. His236 is part of the loop that accommodates $Ca2^{+}$, which has been shown to be essential for the maximal catalytic efficiency and factor VIIIa binding of factor IXa.[11] As we recently demonstrated (factor IX_{Delhi}), this substitution probably affects the conformation of the surface-loop of the catalytic domain and may alter Ca^{2+} accommodation, thereby contributing to the severity of the disease due to hampered post-activational activity of factor IXa.[12] An A→G transition at nucleotide position 17689 in exon 5 leading to the amino acid change Asn92 to Asp in the EGF2 domain was observed in two cases (HBF 6 and HBF 8). Previously Nishimura H *et al.*[13] reported that substitution of Asn92 by His (factor $IX_{Fukuoka}$) leads to moderate disease whereas Chang *et al.*[14] reported that replacement of this residue by Ala reduces the clotting activity of factor IX to less than 0.5%, thus emphasizing the importance of this residue. It has also been proposed that residues 89-94 in the EGF2 domain are involved in binding to factor VIIIa.[14] The substitution of the uncharged polar Asn92 by the negatively charged Asp would therefore have an untoward effect on binding of factor IXa with factor VIIIa leading to severe disease. Indeed, one of these patients (HBF 8) died of severe intracranial hemorrhage. Further computational or experimental investigations are needed to establish the importance of the replacement of Asn92 by different residues (His, Ala, Asp), leading to different effects on factor IX activity.

In three patients belonging to HBF 4, HBF 11, HBF 5, a G→A transition was observed at the nucleotide positions 6428, 17786 and 31170 in exons 2, 5 and 8, respectively. In all three cases this transition led to the substitution of Cys by Tyr, at residues 18, 124 and 350, respectively. As is known, all the Cys residues in factor IX protein, except for two in the leader sequence, are involved in disulfide bond formation. These disulfide bridges establish the functional conformation of factor IX protein and are conserved in all vitamin K-dependent serine proteases. Cys18 forms a disulfide bond with Cys23 in the Gla domain. Loss of this bridge in the Gla domain is likely to distort its 3-dimensional structure such that surface presentation of Gla residues for calcium binding is no longer effective.[15] Similarly Cys124 is involved in formation of a disulfide bridge with Cys111 in the EGF2 domain and Cys350 with Cys336 in the catalytic domain. Disruption of these covalent bonds disturbs the conformation of the respective domains, resulting in a severe phenotype.

A C→T transition was observed in HBF 14, HBF 17, HBF 18 at nucleotide position 6460 in exon 2 and in HBF 13 at position 30863 in exon 8. These transitions led to nonsense mutations at residues Arg29 and Arg248, respectively. The outcome of a nonsense mutation is expected to be a truncated protein molecule and hence these mutations are particularly detrimental and are associated with severe disease. Both these mutations were observed in CpG sites, which are considered to be hypermutable. A C→T transition at nucleotide position 6364 in exon 2 was observed in HBF 12. This leads to the replacement of Arg -4 to Trp in the pre-propeptide, such that this

Table 2. Molecular events in the factor IX gene in North Indian families with hemophilia B.

Family No.	*Individual* Information*	*Genotype*	*Mutation*	*At CpG*	*Location*	*Codon change*	*Evolutionary° Conservation*	*Referred (No. of times)*
HBF 1	Propositus Mother		Nil					
HBF 2	Propositus Mother	C AC	30105A→C	No	Exon 7 Catalytic domain	ACT→CCT Thr218Pro	P	This study
HBF 3	Propositus Mother		Nil					
HBF 4	Propositus Mother	A GA	6428G→A	No	Exon 2 Gla domain	TGT→TAT Cys18Tyr	G	Two
HBF 5	Propositus Mother Sister	A GA GA	31170G→A	No	Exon 8 Catalytic domain	TGT→TAT Cys350Tyr	G	Seven
HBF 6	Propositus Mother	G AG	17689A→G	No	Exon 5 EGF-2 domain	AAT→GAT Asn92Asp	G	Two
HBF 7	Propositus Mother	G TT	30925T→G	No	Exon 8 Catalytic domain	CAT→CAG His268Gln	P	Two
HBF 8	Propositus Mother Maternal Aunt	G AG AG	17689A→G	No	Exon 5 EGF-2 domain	AAT→GAT Asn92Asp	G	Two
HBF 9	Propositus Mother	A TA	30829T→A	No	Exon 8 Catalytic domain	CAT→CAA His236Gln	N	This study
HBF 10	Propositus Mother		Nil					
HBF 11	Propositus Mother	A GA	17786G→A	No	Exon 5 EGF-2 domain	TGT→TAT Cys124Tyr	G	Four
HBF 12	Propositus Mother	T CT	6364C→T	Yes	Exon 2 Pre-propeptide	CGG→TGG Arg-4Trp	P	Forty-five
HBF 13	Propositus Mother	T CT	30863C→T	Yes	Exon 8 Catalytic domain	CGA→TGA Arg248Ter		Fifty-one
HBF 14	Propositus Mother	T CT	6460C→T	Yes	Exon 2 Gla domain	CGA→TGA Arg29Ter		Fifty-two
HBF 15	Propositus Mother	A GA	17798G→A	No	Intron 5	Donor splice site		Seven
HBF 16	Propositus Mother		31042-31051 del AAGTGGCTGG		Exon 8 Catalytic domain	Frame shift; Premature stop in exon 8		This study
HBF 17	Propositus Mother Maternal Aunt Sister	T CT CT CT	6460C→T	Yes	Exon 2 Gla domain	CGA→TGA Arg29Ter		Fifty-two
HBF 18	Propositus Mother Maternal Grandmother	T CT CT	6460C→T	Yes	Exon 2 Gla domain	CGA→TGA Arg29Ter		Fifty-two

**All the patients suffer from severe hemophilia B (clotting activity <1%); °evolutionary conservation:[20] G: generically conserved; P: partially generic; N: non-conserved.*

Table 3. Polymorphisms observed in factor IX gene in North Indian hemophilia B patients.

Family	*Individual information*	*Genotype*	*Polymorphism*	*Location*	*Amino acid variation*
HBF 4	Propositus	C	31231T→C	Exon8	Val370Val
	Mother	TC		Catalytic domain	
HBF 8	Propositus	C	30053T→C	Exon 7	Gly200Gly
	Mother	CT		Catalytic domain	
HBF 11	Propositus	C	31060C→A	Exon 8	Val313Val
	Mother	CA		Catalytic domain	
HBF 14	Propositus	A	17673A→G	Exon 5	Val86Val
	Mother	GA		EGF-2	

missense mutation results in the secretion of factor IX protein with the propeptide covalently attached. Although γ carboxylation appears to occur, the presence of intact propeptide inhibits the normal calcium-dependent conformation of the Gla domain, explaining the severity of the hemophilia B.[16]

A T→G transversion was observed in exon 8 at nucleotide position 30925 in the propositus of HBF7. This leads to the replacement of His268 by Gln in the catalytic domain. We confirmed the presence of this mutation by sequencing a second independent PCR product from the propositus. We were not able to detect this mutation in the patient's family. Sequencing of the appropriate PCR product obtained from mother's and also sister's DNA revealed only the normal factor IX sequence. This indicates that this T→G transversion might have arisen as a result of maternal somatic and/ or germline mosaicism or as a *de novo* event in the patient. His268 lies in the vicinity of Asp269, part of the catalytic triad. The introduction of an uncharged, polar Gln at 268 could perhaps disturb the orientation of Asp269 in the catalytic triad thereby severely affecting the catalytic function of this protease.

We identified a G→A transition at nucleotide position 17798 in the obligatory GT sequence of the donor splice site at the 5' end of intron 5 of the factor IX gene in HBF 15. This donor mutation is found at the base that is a target for virtually complete loss of normal splicing as predicted by Ketterling *et al.*[17] thus leading to the disease. A novel deletion of ten nucleotides (31042-31051 del AAGTGGCTGG in exon 8) leading to frame shift as well as deletion of three residues Ser308, Gly309, Trp310 in the catalytic domain was identified in HBF 16. This frame shift leads to a premature stop in exon 8. The patient's mother was heterozygous for this deletion. A deletion of three residues followed by termination of the reading frame would definitely be detrimental, causing the disease. In three families (HBF 1, HBF 3 and HBF 10) we could not detect any mutations in the exonic regions or splice junctions. It was possible that the change in the three remaining patients could have been in either the promoter region at the 5' end of the gene or in the Poly A addition site at the 3' end. However, PCR amplification and sequencing of these regions excluded this possibility in all three cases. Having excluded all the regions of the gene it is likely that the mutation in these three patients form cryptic splice sites within an intron of the factor IX gene, interfering with the normal splicing pattern,[15] or that these families might have pathological translocations, duplications or inversions in the factor IX gene leading to the disease. However, further investigation is needed in these cases to explore these possibilities.

We report that out of the 14 polymorphic positions reported in the Haemophilia B Mutation database in the regions of factor IX that we sequenced, none was polymorphic in the 63 individuals we analyzed. In addition to the mutations described above, four novel single base mismatches (31231 T→C, 30053T→C, 31060C→A, 17673A→G) were detected in some of the subjects we analyzed. All these mismatches led to synonymous changes within the coding sequence (Table 3). Another finding was that all our patients, except HBF 2, have the Thr type of dimorphism (Malmo polymorphism), i.e., they contained a Thr at residue 148 of the activation peptide of factor IX.[18] This agrees with previous findings that most of the Asian population contains the Thr type of factor IX.[19] A T→C transition at nucleotide position 32847 is reported to be a mutation causing the disease in the Haemophilia B Mutation database, whereas we detected the C variant in all samples, including those from healthy subjects.

In conclusion, our data confirm the remarkable heterogeneity of the mutational spectrum giving rise to

severe hemophilia B (although the clinical expression can also be heterogeneous. It seems possible to establish genotype/phenotype relationships. Our results indicate that point mutations tend to predominate in the as yet studied hemophilia B families. This is probably the first report of mutations in the factor IX gene in North Indian hemophilia B patients of Indo-Aryan lineage. Since India is a land of vast genetic resource, this study can be taken as a lead to derive a complete spectrum of molecular events in hemophilia B patients of this population.

AM gave a substantial contribution to the conception and design of the study and analysis and interpretation of the data. SC was involved in analysis and interpretation of the data. MK and MC were involved in collecting the samples and their clinical evaluation. DB conceived and supervised the study. All authors were involved in drafting the article. We thank Prof. S. K. Brahmachari, Director, IGIB for his support. We are grateful to the patients and their family members for their generous collaboration.

This study was done under the CSIR Task Force project 'Predictive medicine using repeat and single nucleotide polymorphisms' (CMM0016). AM and SC are grateful to CSIR, Government of India for providing the pre-doctoral fellowships. The authors reported no conflict of interest.

Manuscript received July 27, 2004. Accepted October 7, 2004.

References

1. HaemophiliaB Mutation database: (Version 12, 2003): http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html
2. Mukherjee S, Mukhopadhyay A, Banerjee D, Chandak GR, Ray K. Molecular pathology of haemophilia B: identification of five novel mutations including a LINE 1 insertion in Indian patients. Haemophilia 2004;10:259-63.
3. Espinos C, Casana P, Haya S, Cid AR, Aznar JA. Molecular analyses in hemophilia B families: identification of six new mutations in the factor IX gene. Haematologica 2003; 88:235-6.
4. Onay UV, Kavakli K, Kilinc Y, Gurgey A, Aktuglu G, Kemahli S, et al. Molecular pathology of haemophilia B in Turkish patients: identification of a large deletion and 33 independent point mutations. Br J Haematol 2003;120:656-9.
5. Bowen DJ. Haemophilia A and haemophilia B; molecular insights. Mol Pathol 2002;55:127-44.
6. Bentley AK, Rees DJ, Rizza C, Brownlee GG. Defective propeptide processing of blood clotting factor IX caused by mutation of arginine to glutamine at position -4. Cell 1986;45:343-8.
7. Jorgensen MJ, Cantor AB, Furie BC, Brown CL, Shoemaker CB, Furie B. Recognition site directing vitamin K-dependent γ-carboxylation resides on the propeptide of factor IX. Cell 1987; 48:185-91.
8. Cheung WF, Hamaguchi N, Smith KJ, Stafford DW. The binding of human factor IX to endothelial cells is mediated by residues 3-11. J Biol Chem 1992;267: 20529-31.
9. Sambrook J, Russel DW. Molecular cloning: a laboratory manual. 3rd edition. New York: Cold Spring Harbor Laboratory Press. 2001.
10. Yoshitake S, Schach BG, Foster DC, Davie EW, Kurachi K. Nucleotide sequence of the gene for human factor IX (anti-hemophilic factor B). Biochemistry 1985;24:3736-50.
11. Mathur A, Zhong D, Sabharwal AK, Smith KJ, Bajaj SP. Interaction of factor IXa with factor VIIIa. Effects of protease domain $Ca2^{+}$ binding site, proteolysis in the autolysis loop, phospholipid, and factor X. J Biol Chem 1997;272:23418-26.
12. Mahajan A, Sharma A, Chavali S, Kabra M, Chowdhury MR, Srinivasan N, et al. Novel missense mutation in the coagulation factor IX catalytic domain associated with severe Haemophilia B- Factor IX_{Delhi}. Haemophilia 2004;10:550-2.
13. Nishimura H, Takeya H, Miyata T, Suehiro K, Okamura T, Niho Y, et al. Factor IX Fukuoka. Substitution of Asn92 by His in the second epidermal growth factor-like domain results in defective interaction with factors VIIa/X. J Biol Chem 1993; 268:24041-6.
14. Chang YJ, Wu HL, Hamaguchi N, Hsu YC, Lin SW. Identification of functionally important residues of epidermal growth factor-2 domain of factor IX by alanine-scanning mutagenesis. Residues Asn (89)-Gly (93) are critical for binding factor VIIIa. J Biol Chem 2002;277:25393-9.
15. Winship PR, Dragon AC. Identification of haemophilia B with mutations in two calcium binding domains of factor IX: importance of a -OH Asp64 Asn change. Br J Haematol 1991;77:102-9.
16. Thompson AR. Molecular biology of hemophilias. In: Coller BS, eds. Progress in hemostasis and thrombosis. Pennsylvania: WB Saunders Company. 1990.
17. Ketterling RP, Drost JB, Scaringe WA, Liao DZ, Liu JZ, Kasper CK, et al. Reported in vivo splice-site mutations in the factor IX gene: severity of splicing defects and a hypothesis for predicting deleterious splice donor mutations. Hum Mutat 1999;13:221-31.
18. McGraw RA, Davis LM, Noyes CM, Lundblad RL, Roberts HR, Graham JB, et al. Evidence for a prevalent dimorphism in the activation peptide of human coagulation factor IX. Proc Natl Acad Sci USA 1985; 82:2847-51.
19. Smith KJ, Thompson AR, McMullen BA, Frazier D, Lin SW, Stafford D, et al. Carrier testing in hemophilia B with an immunoassay which distinguishes a prevalent factor IX dimorphism. Blood 1987;70:1006-13.
20. Bottema CD, Ketterling RP, Ii S, Yoon HS, Phillips JA, Sommer SS. Missense mutations and evolutionary conservation of aminoacids: evidence that many of the aminoacids in factor IX function as "spacer" elements. Am J Hum Genet 1991;49:820-38.