Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 10095–10099, December 1989
Medical Sciences

Canine hemophilia B resulting from a point mutation with unusual consequences

(factor IX/enzymatic amplification)

JAMES P. EVANS*†, KENNETH M. BRINKHOUS‡§, GARY D. BRAYER¶, HOWARD M. REISNER‡§, AND KATHERINE A. HIGH*‡§‖

Departments of *Medicine and ‡Pathology, and §Center for Thrombosis and Hemostasis at The University of North Carolina, Chapel Hill, NC 27599; and ¶Department of Biochemistry, University of British Columbia, Vancouver, BC, V6T 1W5 Canada

Contributed by Kenneth M. Brinkhous, September 11, 1989

ABSTRACT **We have used the polymerase chain reaction to amplify the entire coding region of canine factor IX from a hemophilia B animal. When the sequence was compared to that which codes for normal canine factor IX, a single missense mutation was identified. This mutation (G → A at nucleotide 1477) results in the substitution of glutamic acid for glycine-379 in the catalytic domain of the molecule. The mutation creates a new restriction site that allowed confirmation of the abnormal sequence in both hemophilic and carrier animals. Amino acid 379 in canine factor IX corresponds to position 381 in human factor IX, a location at which no human mutations have been described. Moreover, it occurs at one of the few amino acids that have been rigorously conserved among the trypsin-like serine proteases throughout evolution. The mutation responsible for canine hemophilia B results in a complete lack of circulating factor IX in the affected animals. As it is unusual for a missense mutation to result in a complete absence of protein product, structural modeling of the mutant and normal proteins was pursued. These studies suggest that the observed mutation would have major adverse effects on the tertiary structure of the aberrant factor IX molecule. The elucidation of this mutation sheds light on structure–function relationships in factor IX and should facilitate future experiments directed toward gene therapy of this disease.**

Factor IX is a 56-kDa plasma glycoprotein that is required for blood coagulation (1). Factor IX circulates as a zymogen; to become active in coagulation, factor IX must undergo limited proteolytic cleavage by factor XIa or factor VIIa-tissue factor. This proteolytic step cleaves two peptide bonds to generate a short amino acid segment termed the activation peptide and unblocks the active site of the enzyme. This activated form, designated factor IXa, in the presence of its cofactor factor VIIIa and calcium and phospholipids, cleaves factor X to generate factor Xa, the next step in the coagulation cascade. Factor IX belongs to the serine protease superfamily, a class of enzymes that includes trypsins as well as coagulation and complement proteins, among others (2). These enzymes arose early in evolution and are present in organisms from bacteria to man. Through evolution, they have acquired a variety of specialized functions. The serine proteases of the coagulation cascade possess, in addition to a trypsin-like active site, domains that are thought to confer cofactor and receptor specificity.

Analysis of the genomic organization of the human factor IX gene has provided insight into the structure of the protein. The gene contains eight exons that can be correlated in a rough way with the functional domains of the protein (3). Exon I codes for the signal sequence, exon II codes for the γ-carboxyglutamic acid-rich region, exon III codes for the short aromatic acid stack, exons IV and V code for the epidermal growth factor-like domains, exon VI codes for the activation peptide, and exons VII and VIII code for the catalytic domain and the long 3′ untranslated region, respectively. Although genomic structure has been reported only for human factor IX, a cDNA sequence is available for canine factor IX (4), and the amino acid sequence is available for bovine factor IX (5). Comparison of sequences among these species reveals >90% conservation of residues in the γ-carboxyglutamic acid-rich region and in the heavy chain and a much lower degree of conservation in the activation peptide.

An absence of functional factor IX results in the sex-linked bleeding disorder hemophilia B. Current treatment of hemophilia B requires the infusion of factor IX-rich plasma concentrates during bleeding episodes. Such treatments must be repeated with each episode of bleeding and carry a risk of contamination by infectious agents. The possibility of treating hemophilia B by gene therapy has been proposed; one approach would involve the introduction of a gene encoding factor IX into the somatic cells of an affected individual. Hemophilia B is an attractive target for gene transfer therapy since neither precise regulation of factor IX levels nor tissue-specific expression is required. A further advantage is the existence of an animal model of hemophilia B (6, 7), a potentially valuable resource in trials of gene therapy (8, 9). To maximize the utility of canine hemophilia B in such studies, it is desirable to know the causative mutation at the molecular level. In this report we describe the isolation and characterization of the cDNA encoding mutant (hemophilic) canine factor IX. The molecular defect responsible for this strain of canine hemophilia B is a missense mutation in a highly conserved residue of the catalytic domain; it results in a complete lack of detectable protein in the circulation of the affected animals.

MATERIALS AND METHODS

Canines. Normal, carrier, and factor IX-deficient canines are maintained at the Francis Owen Blood Research Laboratory, The University of North Carolina at Chapel Hill.

Determination of Factor IX Antigenic Levels in Hemophilic Canines. Quantitation of circulating factor IX levels was by ELISA (10). All samples were assayed in duplicate. In addition, a radioimmunoassay utilizing an affinity-purified

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: PCR, polymerase chain reaction.

†Present address: Department of Genetics, University of Washington, Seattle, WA 98195.
‖To whom reprint requests should be addressed at: Division of Hematology, Department of Medicine, CB#7035, The University of North Carolina, Chapel Hill, NC 27599.

rabbit anti-canine factor IX antibody was performed by Arthur Thompson (Puget Sound Blood Center, Seattle).

Isolation of Hepatic DNA and RNA: Southern and Northern Blot Analysis. High molecular weight DNA was extracted from hemophilic canine liver by a modification of the method of Blin and Stafford (11). Total RNA was isolated from frozen (−70°C) dog liver using the guanidinium thiocyanate method of Chirgwin *et al.* (12). For Southern (13) and Northern (14) blot analysis, a ^{32}P-labeled human factor IX cDNA was used as a probe (15).

Reverse Transcription and DNA Amplification. Synthesis of single-stranded complementary DNA from a template of total hepatic RNA was carried out by the method of Gubler and Hoffman (16) with several modifications (4). $(dT)_{12-18}$ was used as a primer for reverse transcription.

Amplification of factor IX reverse transcripts was achieved by subjecting the single-stranded cDNA to the polymerase chain reaction (PCR) (17). Priming oligonucleotides were based on the sequence of normal canine factor IX (4). A typical experiment involved reverse transcription of 50 μg of total hepatic RNA. The reaction mixture was extracted with phenol and chloroform and precipitated overnight in ethanol at −20°C. Twelve percent of the total was centrifuged, washed, and dissolved in 10 μl of 10× PCR buffer (670 mM Tris·HCl, pH 8.8/67 mM EDTA/100 mM 2-mercaptoethanol/166 mM $(NH_4)_2SO_4$/67 mM $MgCl_2$). The final volume of the reaction was 100 μl and contained 10% (vol/vol) dimethyl sulfoxide, 1.5 mM deoxynucleoside triphosphates, oligonucleotide primers (each at a concentration of 0.5 μM), and 2.5 units of thermostable polymerase (Cetus). Amplification was carried out manually with an initial denaturing step at 94°C for 7 min. Annealing was at 55°C for 3 min, extension at 70°C for 4 min, and denaturation at 94°C for 3 min. The reactions were subjected to 35 cycles of amplification after which agarose electrophoresis and electroelution were used to isolate the appropriate fragments. When genomic DNA was used as starting material, a 133-base-pair (bp) fragment containing the mutation was amplified. Priming oligonucleotides were constructed based on the normal canine sequence (4) (GenBank accession no. M21757). *Dde* I (New England BioLabs) was used according to manufacturer's recommendations.

Subcloning and DNA Sequencing. The termini of the amplified DNA were phosphorylated with T4 nucleotide kinase (New England BioLabs) and the fragments were subcloned into dephosphorylated M13 vectors. Sequencing was by the method of Sanger *et al.* (18), and in certain regions oligonucleotides were constructed for use as primers in the sequencing reactions. The entire coding region of factor IX as well as 149 bp of 3′ untranslated material were sequenced in both orientations. When discrepancies arose between the normal canine factor IX sequence (4) and that obtained from the hemophilic animal, the area in question was subjected to two additional independent amplifications from newly synthesized cDNA to eliminate the possibility of a polymerase error masquerading as a mutation.

Amino Acid Sequence Alignment and Structural Modeling. Alignment of the primary structure of the protease domain of canine factor IX with those of human factor IX, trypsin, chymotrypsin, elastase, kallikrein, and *Streptomyces griseus* protease A was based on sequence identity and structural topological equivalence (19–21). Modeling of human factor IX has been described (21) and was based primarily upon the known high-resolution structure of bovine pancreatic trypsin (22).

RESULTS

Quantitation of Factor IX Levels in Canines. The ELISA which was employed relies on a monoclonal capture antibody mapped to the first epidermal growth factor-like domain of human factor IX and a polyclonal antibody raised to human factor IX. Due to the extensive conservation that exists between human and canine factor IX (4), this assay is able to detect the canine protein specifically at plasma dilutions of 1:320. When plasma from factor IX-deficient canines was analyzed in this assay, no factor IX was detected (data not shown). These results were confirmed and extended by a radioimmunoassay that utilized a polyclonal rabbit anti-canine factor IX antiserum. This assay is able to detect canine factor IX at plasma concentrations of <0.3%. When coded samples were assayed by this method, again no factor IX was detectable.

Restriction Mapping of the Factor IX Gene and Analysis of the Transcript. Southern blots of normal and hemophilic canine genomic DNA probed with a factor IX cDNA were identical (data not shown), indicating that canine hemophilia B does not result from a grossly detectable abnormality of the factor IX gene. Total hepatic RNA from normal and hemophilic canines was subjected to Northern blot analysis, as shown in Fig. 1. Hemophilic animals produce a factor IX transcript; it is similar in size (approximately 3 kilobases) to the normal and is present at approximately the same steady-state levels.

Nucleotide Sequence Analysis of Hemophilic Canine Factor IX. The data from Southern and Northern blot analyses indicated that a point mutation most likely accounted for the defect in canine hemophilia B. Therefore, total hepatic RNA from a factor IX-deficient animal was subjected to reverse transcription, and the coding region of canine factor IX was amplified by the PCR technique. Amplification utilized pairs of oligonucleotide primers flanking the entire coding region of canine factor IX, including 5′ and 3′ untranslated sequence. The ends of the amplified fragments overlapped so that any mutation present in a region where a primer annealed would be detected. Sequencing revealed a missense mutation at nucleotide 1477 (G → A) in the region encoding the catalytic domain of the hemophilic factor IX gene (Fig. 2). This mutation changes a glycine to a glutamic acid in the heavy chain of the molecule (amino acid 379 in the canine). It corresponds to amino acid 381 in human factor IX and position 211 in the chymotrypsinogen numbering system of the trypsin-like serine proteases. This is the only difference in the sequence of normal and hemophilic canine factor IX, throughout the coding region and throughout 149 bases of untranslated sequence. This region was subjected to three independent amplifications of newly synthesized cDNA and identical results were obtained on sequencing in each case. The mutation that we describe creates a new restriction enzyme recognition site for *Dde* I. Thus it was possible to

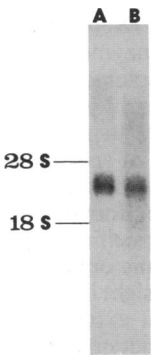

FIG. 1. Northern blot analysis of total RNA from hemophilic (lane A) and normal (lane B) canine hepatocytes. Each lane was loaded with 10 μg of total liver RNA. Blots were probed with a 1500-bp nick-translated fragment of normal canine factor IX (4). Positions of 28S and 18S rRNA are indicated.

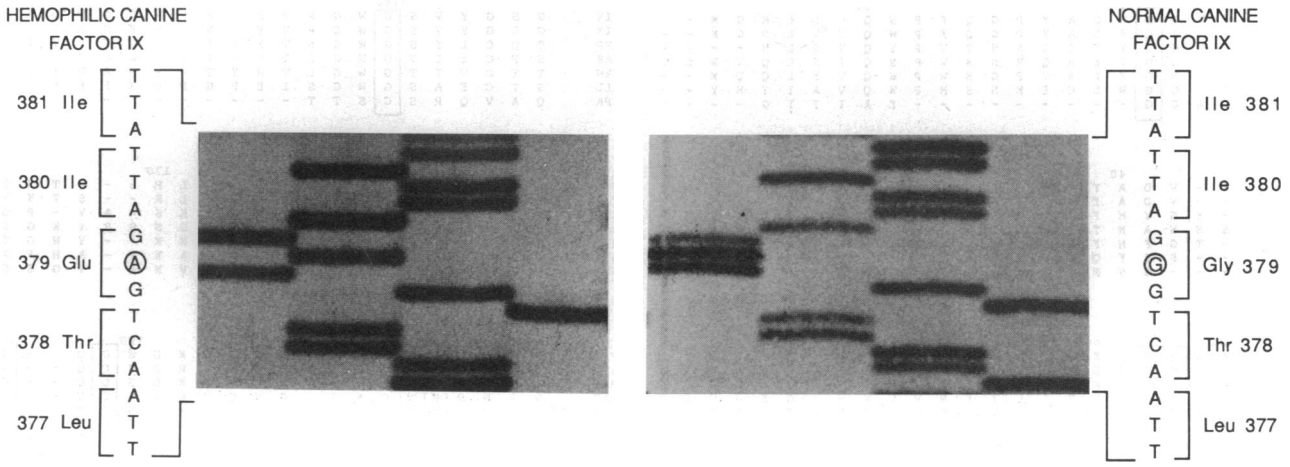

FIG. 2. Nucleotide sequences of hemophilic and normal canine factor IX along with the respective translation products between amino acid positions 377 and 381. The change from a guanine to an adenine in the hemophilic animal results in the substitution of glutamic acid for glycine at position 379.

determine whether other animals within the colony had the same mutation. A 133-bp fragment surrounding the mutation was amplified by PCR from genomic DNA and digested with *Dde* I. The results, shown in Fig. 3, demonstrate that the four normal animals tested do not contain the mutation and that the carrier and hemophilic animals do (demonstrated by the presence of the 79-bp and 54-bp fragments).

Protein Alignment. The catalytic domain of the trypsin-like serine proteases has been highly conserved throughout evolution. Amino acid alignment of the members of this family of proteases demonstrates that certain amino acids have remained invariant from bacteria to humans. Fig. 4 shows the amino acid alignment of the heavy chain of canine factor IX with the corresponding region of several other serine proteases (human factor IX, trypsin, elastase, chymotrypsin, kallikrein, and *Streptomyces griseus* protease A). As can be seen, the mutation in canine factor IX affects one of these highly conserved amino acids, substituting a negatively charged residue (glutamic acid) for an amino acid that is uncharged and possesses no side chain (glycine).

Structural Modeling. The substitution of a glutamic acid for a glycine within the catalytic domain of canine factor IX would be likely to have profound structural consequences. To investigate this possibility further, a structural model of the mutant protein was constructed. Fig. 5*A* shows a structural model of the heavy chain of human factor IX. This should be representative of the structure of the canine molecule, since, for the amino acids illustrated, there is complete conservation in identity between human and canine factor IX. Fig. 5*B* depicts this region after substitution of a glutamic acid for glycine at position 381, the mutation found in the hemophilic canine. This model suggests that substitution of a glutamic acid at position 211 leads to prohibitive spatial conflicts with several residues (Leu-162, Phe-181, His-199, and Tyr-228) and alters the hydrophobic character of this closely packed and internal region of factor IX.

DISCUSSION

The characterization of mutations that give rise to hemophilia B in humans has provided insight into critical functional regions of the factor IX molecule. Since the heavy chains of human and canine factor IX show remarkable identity (4), our description of the mutation responsible for canine hemophilia B provides similar insight into structure–function relationships of the human molecule. There has been no previous report of a human mutation at the site corresponding to the one we report here in canine hemophilia (position 379 in canine and position 381 in human factor IX). This mutation occurs at a residue that has remained invariant in serine proteases from bacteria to humans. The profound effect of the mutation on factor IX levels suggests that the glycine at position 379 (position 211 in the chymotrypsinogen numbering system) plays a critical role in the integrity of the trypsin-like serine proteases in general.

Hemophilia B has been classified traditionally as either cross-reacting-material-positive in which antigenically detectable protein is present or cross-reactive-material-negative in which protein is not detected. The hemophilic canines were found to have undetectable levels of factor IX by two independent assays. This represents an unusual situation among factor IX mutations, since in those characterized to date, missense mutations have always resulted in detectable levels of factor IX (23). Indeed, it is unusual in general for a missense mutation to result in an absence of

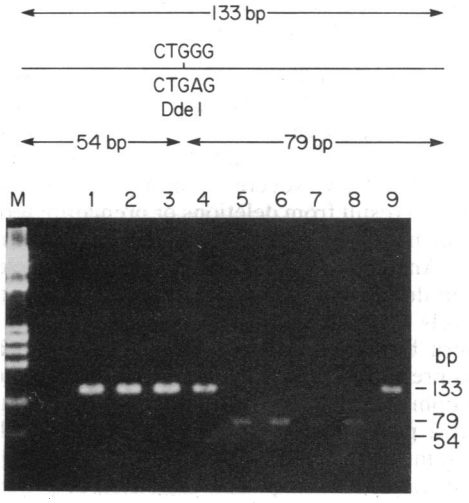

FIG. 3. *Dde* I analysis of amplified fragment coding for the region of the heavy chain containing the mutation. (*Upper*) *Dde* I recognizes the nucleotide sequence CTNAG, where N is any nucleotide. The G → A transition in the hemophilic allele results in the creation of a new restriction site for *Dde* I. (*Lower*) Lanes: 1–4, PCR-amplified DNAs from four unrelated normal canines were not cleaved by *Dde* I and thus do not contain the mutation; 5–8, PCR-amplified DNAs from hemophilic animals have only the 79- and 54-bp fragments; 9, PCR-amplified DNA from a carrier animal has the 133-bp fragment from the normal allele and the 79- and 54-bp fragments from the mutant allele; M, marker lane containing ϕX174 digested with *Hae* III.

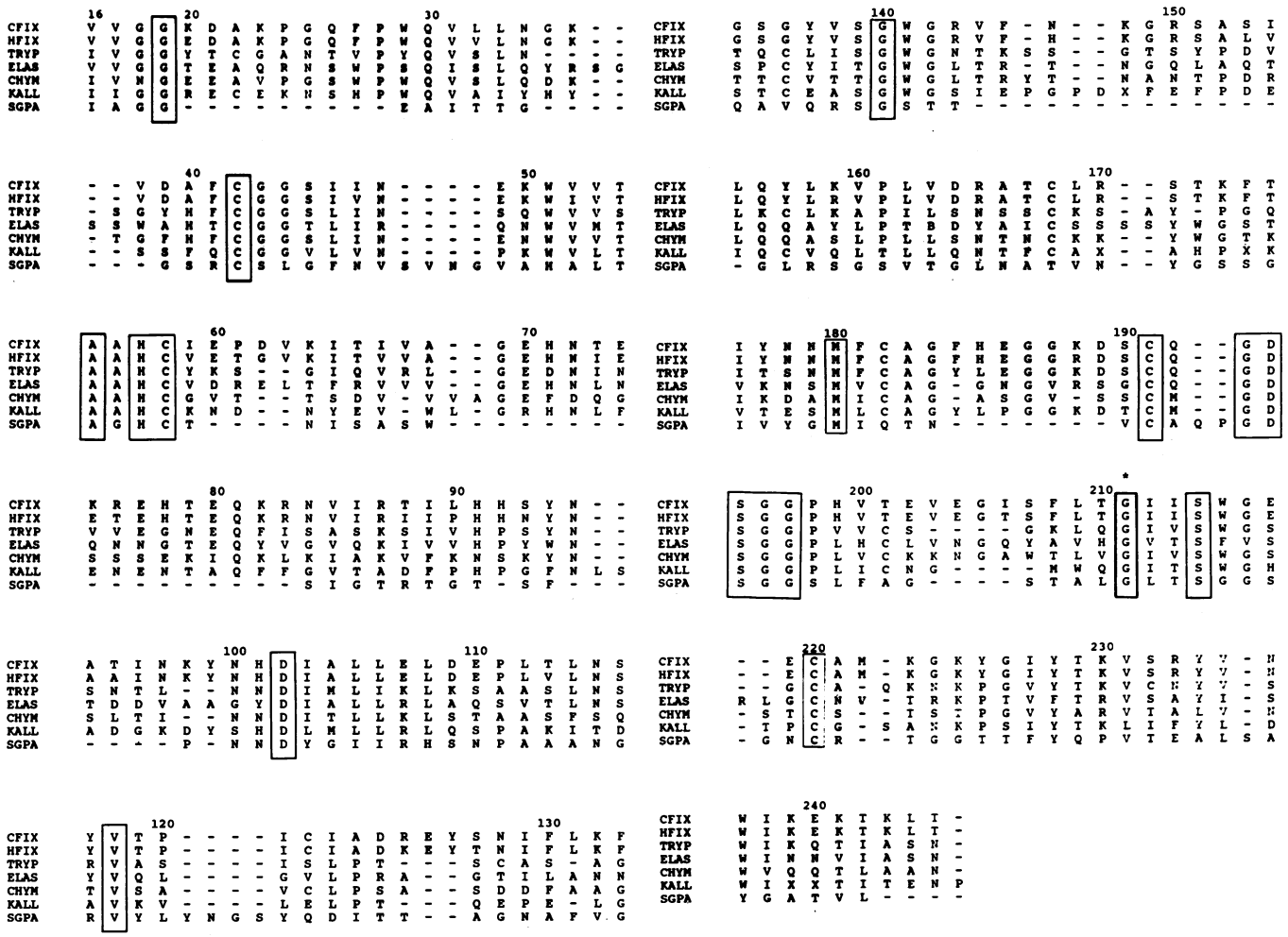

FIG. 4. Amino acid sequence alignment of the protease domain of canine factor IX (CFIX) with human factor IX (HFIX), bovine trypsin (TRYP), porcine elastase (ELAS), bovine chymotrypsin (CHYM), porcine kallikrein (KALL), and *Streptomyces griseus* protease A (SGPA). Conserved amino acids are boxed. The position of the mutation in canine hemophilia is denoted by an asterisk. The numbering system is that of chymotrypsinogen A, and deletions are indicated by dashes.

detectable translated protein. For example, >400 missense mutations have been described in the globins (representing 25% of all possible mutations) (24) and only 2 of these (25, 26) result in stable mRNA but a lack of detectable protein in the circulation. Similarly, severe α_1-antitrypsin deficiency has been found to result from deletions or premature termination codons but has thus far not been attributed to a missense mutation. Among >80 missense mutations analyzed in α_1-antitrypsin deficiency, all result in the production of appreciable levels of protein (27).

Northern blot analysis (Fig. 1) indicated that factor IX mRNA is present in similar steady-state levels in the normal and hemophilic dogs, suggesting that failure of expression occurs at a posttranscriptional level. Several explanations could account for the absence of detectable protein, including decreased translation, impaired secretion, or production of an unstable polypeptide. Structural modeling of factor IX suggests that folding of the mutant protein is likely to be severely disrupted and thus offers a possible explanation for the absence of factor IX. The tertiary structure of factor IX has not been determined using x-ray diffraction techniques; however, firm structural data exist for a number of trypsin-like serine proteases, and there is good reason to believe that the degree of sequence identity that exists between the clotting proteins and other members of this enzyme family translates into a high degree of structural homology (20, 21, 28). Examination of the factor IX model predicts that the presence of a glutamic acid side chain at position 211 would result in significant steric hindrance that in turn would cause improper folding of the molecule and structural instability. As Fig. 5 shows, replacement of glycine with a glutamic acid residue results in prohibitively close contacts with Leu-162, Phe-181, His-199, and Tyr-228, all of which are tightly packed together in this central and internal region of the enzyme. Furthermore, the charged character of the glutamic acid side chain is incompatible with the hydrophobic nature of these surrounding residues.

The fact that other hemophilic animals demonstrate the same mutation does not eliminate the possibility that the mutation we describe is a tightly linked polymorphism; however, it does demonstrate unequivocally that the mutation segregates with the hemophilic phenotype. This, coupled with the structural modeling data and the absence of any other mutations, strongly suggests that the glycine → glutamic acid change is indeed the causative mutation.

The canine model of hemophilia B represents a valuable resource in trials of gene therapy, and the characterization of the defect will facilitate use of the animal model. The lack of antigenically detectable factor IX is an advantage for gene therapy trials, since even low levels of expression of a foreign factor IX gene should be readily distinguished from baseline.

We thank Dr. Arthur Thompson whose canine factor IX radioimmunoassay provided valuable confirmatory data, G. V. Louie for

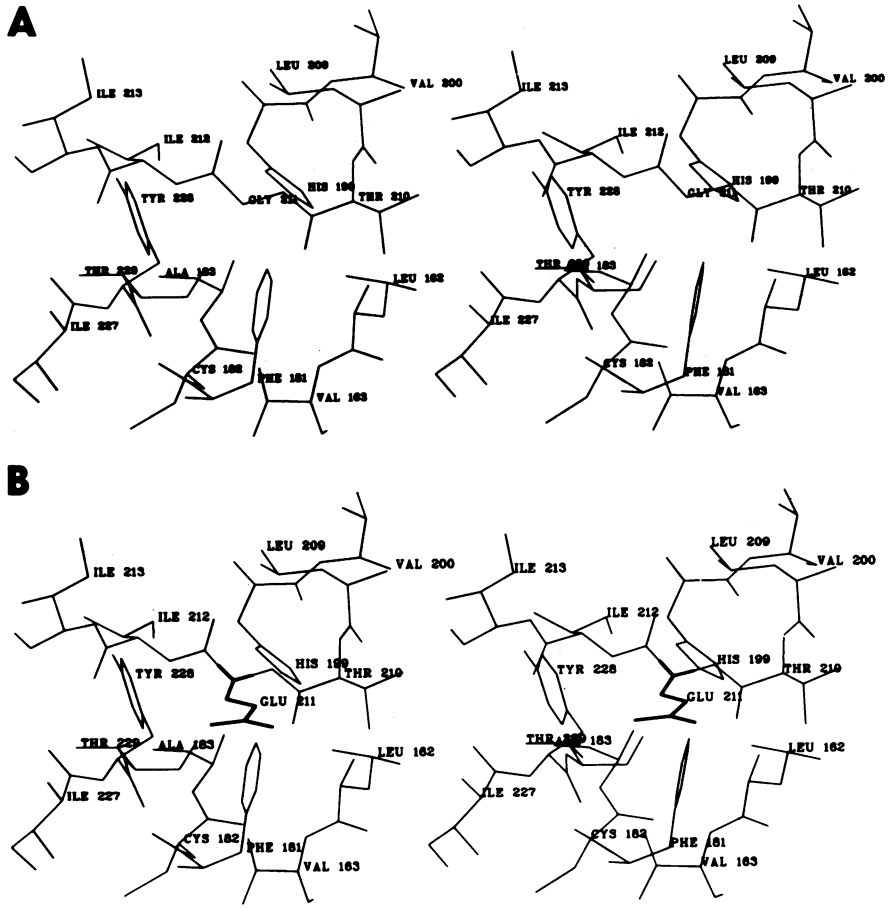

FIG. 5. (*A*) Stereodrawing of human factor IX catalytic domain. (*B*) Stereodrawing of human factor IX with glutamic acid substituted for glycine at position 211.

graphics support, and Sherry Stanfield-Oakley and the staff of the Francis Owen Blood Research Laboratory for valuable technical assistance. This work was supported by Grants HL01648, HL06350, and HL26309 from the National Institutes of Health. J.P.E. was supported by a Judith Graham Pool research fellowship in hemophilia from the National Hemophilia Foundation and K.A.H. was supported by Clinical Investigator Award K08-HL01922 from the National Heart, Lung and Blood Institute. G.D.B. was supported by a Medical Research Council of Canada operating grant.

1. Furie, B. & Furie, B. C. (1988) *Cell* **53,** 505–518.
2. Neurath, H. (1984) *Science* **224,** 350–357.
3. Yoshitake, S., Schach, B. G., Foster, D. C., Davie, E. W. & Kurachi, K. (1985) *Biochemistry* **24,** 3736–3750.
4. Evans, J. P., Watzke, H. H., Ware, J. L., Stafford, D. W. & High, K. A. (1989) *Blood* **74,** 207–212.
5. Fujikawa, K., Legaz, M. E., Kato, H. & Davie, E. W. (1974) *Biochemistry* **22,** 4508–4516.
6. Brinkhous, K. M., Davis, P. D., Graham, J. B. & Dodds, W. J. (1973) *Blood* **41,** 577–585.
7. Brinkhous, K. M., Hedner, U., Garris, J. B., Diness, V. & Read, M. S. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1382–1386.
8. St. Louis, D. & Verma, I. M. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3150–3154.
9. Palmer, T. D., Thompson, A. R. & Miller, A. D. (1989) *Blood* **73,** 438–445.
10. Denton, P. H., Fowlkes, D. M., Lord, S. T. & Reisner, H. M. (1988) *Blood* **72,** 1407–1411.
11. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3,** 2303–2308.
12. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
13. Southern, E. (1975) *J. Mol. Biol.* **98,** 503–517.
14. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
15. McGraw, R. A., Davis, L. M., Noyes, C. M., Lundblad, R. L., Roberts, H. R., Graham, J. B. & Stafford, D. W. (1980) *Proc. Natl. Acad. Sci. USA* **82,** 2847–2851.
16. Gubler, V. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
17. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
19. Cool, D. E., Edgell, C.-J. S., Louie, G. V., Zoller, M. J., Brayer, G. D. & MacGillivray, R. T. A. (1985) *J. Biol. Chem.* **260,** 13666–13676.
20. James, M., Delbaere, L. & Brayer, G. D. (1978) *Can. J. Biochem.* **56,** 396–402.
21. Geddes, V. A., Le Bonniec, B. F., Louie, G. V., Brayer, G. D., Thompson, A. R. & MacGillivray, R. T. A. (1989) *J. Biol. Chem.* **264,** 4689–4697.
22. Marquart, M., Walter, J., Deisenhofer, J., Bode, W. & Huber, R. (1983) *Acta Crystallogr. Sect. B* **39,** 480–487.
23. Lozier, J. N. & High, K. A., *Hematologic Pathology*, in press.
24. Bunn, H. F. & Forget, B. G. (1986) *Hemoglobin: Molecular, Genetic, and Clinical Aspects* (Saunders, New York), pp. 223–321.
25. Liebhaber, S. A. & Kan, Y. W. (1983) *J. Clin. Invest.* **71,** 461–466.
26. Adams, J. G., Steinberg, M. H., Boxer, L. A., Baehner, R. L., Forget, B. G. & Tsistrakis, G. A. (1979) *J. Biol. Chem.* **254,** 3479–3482.
27. Brantly, M., Nukiwa, T. & Crystal, R. G. (1988) *Am. J. Med.* **84,** Suppl. 6A, 13–31.
28. Madison, E. L., Goldsmith, E. J., Gerard, R. D., Gething, M.-J. H. & Sambrook, J. F. (1989) *Nature (London)* **339,** 721–724.