Factor IX\textsubscript{Alabama}: A Point Mutation in a Clotting Protein Results in Hemophilia B

By Lisa M. Davis, Royal A. McGraw, Jerry L. Ware, Harold R. Roberts, and Darrel W. Stafford

Factor IX\textsubscript{Alabama} is a variant factor IX molecule responsible for a clinically moderate form of hemophilia B. Twenty-five kilobases (kb) of the variant gene, including seven exons coding for the structural protein, were cloned and characterized. The restriction map and the arrangement of coding regions are identical to those of the normal gene. DNA sequence analysis of the coding regions revealed a single base-pair difference between the gene for factor IX\textsubscript{Alabama} and the normal factor IX gene. An adenine to guanine transition in the first nucleotide of exon d causes the substitution of a glycine codon (GGT) for the normal aspartic acid codon (GAT). This point mutation results in a single amino acid substitution at residue 47 of the zymogen and represents the genetic defect in factor IX\textsubscript{Alabama}.

MATERIALS AND METHODS

Patients. Patients from the pedigree with factor IX\textsubscript{Alabama} have been previously described.\textsuperscript{12}\textsuperscript{13} Factor IX was purified from the plasma of affected patients.\textsuperscript{14} After informed consent from patients, DNA was extracted from whole blood or leukocytes from the two affected patients in this kindred list.\textsuperscript{15}

Factor IX\textsubscript{Alabama}. Factor IX\textsubscript{Alabama} was isolated from the patients' plasma as previously described.\textsuperscript{14} The mol wt and determination of ω-carboxyglutamic acid residues were performed essentially as described for factor IX\textsubscript{Chapel Hill}.\textsuperscript{16} β-hydroxyaspartic acid was determined by Dr Walter Kisiel of the University of Washington, Seattle.\textsuperscript{17}

Factor IX\textsubscript{Alabama} genomic library. The recombinant library was constructed by ligating a size-fractionated EcoRI partial digest of this genomic DNA to the phage vector λ gt11.\textsuperscript{18} Conditions for sucrose gradient fractionation, ligation, and λ packaging were as described.\textsuperscript{19}

Factor IX\textsubscript{Alabama} clones. The library was screened for factor IX\textsubscript{Alabama} clones with synthetic oligonucleotide probes or with the full-length factor IX cDNA.\textsuperscript{4} Oligonucleotides were end-labeled as described,\textsuperscript{19} and the full-length cDNA was radiolabeled according to a standard nick-translation protocol.\textsuperscript{20} Alternatively, the full-length cDNA, subcloned into the M13 phage vector mp9, was labeled with α\textsuperscript{32P}dATP in a reaction similar to the chain-terminating deoxy DNA-sequencing reaction except that deoxy nucleotides were omitted. The factor IX\textsubscript{Alabama} clones were subcloned from the λ phage vector into both pUC plasmids and M13 phage vectors to facilitate restriction analysis and DNA sequence determination. Standard procedures for construction of these subclones were followed.\textsuperscript{21}\textsuperscript{23}

DNA sequence analysis. DNA sequence analysis was done according to the method of Biggin and co-workers\textsuperscript{30} by a modification of the Sanger deoxy chain-terminating method\textsuperscript{25} using end-labeled primers.\textsuperscript{20} Primer was either the universal pentadecamer (New England Biolabs) or one of the factor IX-specific synthetic oligonucleotides.\textsuperscript{1} The template was either an M13 single-stranded DNA or a pUC double-stranded DNA subclone.

RESULTS

The mol wt of factor IX\textsubscript{Alabama} was normal. The protein has 12 ω carboxyglutamic acid residues and a β-hydroxy-
aspartic acid at position 64 (data not shown). Factor IX can be activated to the factor IXaβ form of the molecule and possesses 10% of the clotting activity of normal factor IXaβ. In addition, factor IXAlabama is present in normal amounts, as previously demonstrated.12

Two factor IX genomic clones, designated LC-7 and LC-10, were isolated from the factor IXAlabama library. Together they contain 25 kb of the factor IXAlabama gene. These clones include all the sequence coding for the mature structural protein and one-third of the leader peptide. The remainder of the pre-pro sequence was not determined, since the factor IXAlabama Protein was secreted normally, as evidenced by normal plasma antigen concentrations, and had a normal complement of γ carboxyglutamic acid residues. These findings indicated that the pre-pro portion of the gene was normal.

LC-7 was identified with an end-labeled 18-mer oligonucleotide probe, synthesized for the active serine site region of exon h. LC-10 was identified with the uniformly labeled, Klenow-extended, full-length cDNA. The identity of each clone was established initially as follows: a Sau 3A digest of each clone was subcloned into the BamHI site of mp8 and rescreened with the original probe. The DNA sequences of positive plaques were determined and compared with the published cDNA sequences.2,4

**DNA sequence analysis.** Figure 1 shows the strategy used in sequencing factor IXAlabama. In those cases in which sequence of only one strand was obtained, several independently derived templates were used. Because exon d contains the point substitution found to cause the functional defect of factor IXAlabama (discussed below), the DNA sequence of this region was determined in both strands from six independent M13 templates derived from different λ bacteriophage preparations.

The nucleotide sequence for ~4,600 bases of the factor IXAlabama gene was determined. This included all the sequence coding for the mature structural protein, one-third of the leader peptide sequence, all the 3' nontranslated region, and some flanking intron sequence (Table 1).

The sequence throughout the mature coding region of factor IXAlabama is identical to all the normal factor IX clones except at two positions. Both affect the primary amino acid sequence. One difference occurs in the third residue of the activation peptide and has been shown by us to result from a nondeleterious dimorphism that is found in both the normal and hemophilic population.4 The presence of either a guanine or an adenine in the first position of the codon for the third residue of the activation peptide results in either an alanine (GCT) or threonine (ACT) in this position. This corresponds to residue 148 of the zymogen. Factor IXAlabama codes for a threonine in this position.

The other difference affecting the amino acid sequence is not found in any of the normal factor IX sequences and constitutes the abnormality in factor IXAlabama. This difference occurs in the first nucleotide of exon d (Fig 2A). This is the second nucleotide of the triplet coding for the 47th residue of the mature protein. An adenine to guanine transition causes the substitution of a GGT codon for the normal GAT codon. This mutation results in a glycine substitution for aspartic acid at residue 47, preceding the epidermal growth-factor-like region of the protein. Figure 2B shows more clearly the DNA coding sequence of normal factor IX and factor IXAlabama in this region and the corresponding translated amino acid sequence.

Two genomic factor IX sequences5,6 are available for comparison of intron sequences that demonstrate naturally occurring polymorphisms. Anson and colleagues7 reported the presence of a thymine between the cytosine and adenine that occur at 124 and 125 nucleotides 5' to exon b. This thymine is absent in both factor IXAlabama and in the normal sequence, as reported by Yoshitake and co-workers.6 The sequence of one cDNA clone5 and two genomic clones5,6 that include the complete 3' nontranslated region have been reported. There are discrepancies between those sequences presented in this paper and those presented by Yoshitake and co-workers.6

---

Table 1. Summary of Determined F9Alabama Nucleotide Sequence

<table>
<thead>
<tr>
<th>Exon</th>
<th>Exon Length</th>
<th>Nucleotides 5'to Exon</th>
<th>Nucleotides 3'to Exon</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>164</td>
<td>190</td>
<td>188</td>
</tr>
<tr>
<td>c</td>
<td>25</td>
<td>188*</td>
<td>9</td>
</tr>
<tr>
<td>d</td>
<td>114</td>
<td>106</td>
<td>140</td>
</tr>
<tr>
<td>f</td>
<td>203</td>
<td>226</td>
<td>35</td>
</tr>
<tr>
<td>g</td>
<td>115</td>
<td>228</td>
<td>40</td>
</tr>
<tr>
<td>h</td>
<td>1,963†</td>
<td>394</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Totals</td>
<td>2,713</td>
<td>1,250</td>
</tr>
</tbody>
</table>

*This sequence is also listed as nucleotides 3' to exon b and has not been included in the total.
†Includes 1,418 nucleotides at the 3' end that are not translated.
and the one in this report at four positions, all involving single base-pair changes. When the sequence numbers of Yoshitake and co-workers are used as a reference, the cDNA sequence and the genomic sequence show cytosine at 31,663 where factor IXAlama has thymine; similarly, both the cDNA clone and the genomic clones have adenine at position 32,056 where factor IXAlama has guanine. Factor IXAlama, however, our cDNA clone, and the normal genomic sequence are alike at two additional positions, (adenine at 32,063 and cytosine at 32,308) but different from the cDNA sequence reported by Anson and colleagues, which has guanine at both positions.

DISCUSSION

Factor IXAlama is a dysfunctional coagulation protein with a normal mol wt and a normal complement of γ carboxyglutamic acid residues, suggesting that the defect is not due to gross structural alterations of the gene. Our results obtained by both restriction and DNA sequence analysis of the clones confirm that the gross structure of the factor IX gene is intact. DNA sequence analysis, however, reveals a single base-pair substitution in the coding region of factor IXAlama, an adenine to guanine transition at nucleotide 307 in the cDNA sequence (corresponds to nucleotide 10,392 in ref 6) changes a GAT codon to a GGT codon, resulting in a glycine substitution for aspartic acid at position 47 of the protein. The substitution occurs in the amino terminal region of the “connecting peptide” in the light chain of factor IX preceding the epidermal growth factor domain. This single point substitution is the only observed structural difference between factor IXAlama and normal factor IX, suggesting that this is the genetic defect responsible for the decreased clotting activity of factor IXAlama. This substitution is sufficient to result in moderate bleeding in affected members of the factor IXAlama kindred.

DNA sequence of exon a, which codes for two-thirds of the leader peptide, was not determined. The leader peptide is believed to function in transport of the protein from the site of synthesis to its site of carboxylation in the cytoplasm of the hematocyte before secretion into the circulation. Because factor IXAlama is normally secreted and carboxylated, it is very unlikely that a defect exists in the region of exon a. Therefore, absence of the sequence of exon a should not affect the conclusions drawn here regarding the defect responsible for the bleeding tendency in affected patients with factor IXAlama.

The DNA sequence of exon d was confirmed in both strands. The possibility that the substitution might be a cloning artifact is diminished by the fact that the sequence was determined from six independent templates derived from two independent preparations of the λ clone LC-10. In addition, primary amino acid sequence analysis of the isolated factor IXAlama protein confirms the presence of glycine instead of aspartic acid at position 47.

Six other nucleotide differences exist between this variant gene and previously published sequences of normal factor IX genes, but five of these do not affect the structural protein and cannot account for the defect in factor IXAlama. The sixth difference is a naturally occurring dimorphism in the third residue of the activation peptide and again cannot be responsible for the dysfunctional protein because the dimorphism is observed in both normal factor IX and factor IXAlama and in other abnormal factor IX molecules. Together, these facts suggest that a glycine substitution for aspartic acid at residue 47 is responsible for the decreased clotting activity of the factor IXAlama molecule. Aspartic acid is a relatively large, charged residue, usually found on the surface of globular proteins, where it may interact with other proteins or cofactors. Substitution of the small, uncharged glycine may interfere with normal charge interactions occurring in this region of the protein. Alternatively, the substitution may lead to an altered conformation by virtue of the relatively flexible C-N bond of glycine. Whether the functional defect of factor IXAlama results primarily from the loss of charge interactions, an altered conformation, or some other function, remains to be established. Previous work from our laboratory has shown, however, that the binding of activated factor IXAlama with phospholipid is normal, although clotting activity is abnormal. Thus, we assume that the functional abnormality causing the decreased clotting activity of factor IXAlama is due to the mutation at position 47 of the zymogen, although the mechanism of the defect is not yet clear and awaits further study.
FACTOR IXA: A POINT MUTATION

ACKNOWLEDGMENT

We thank Dr John Graham for help in obtaining blood samples, Drs Roger Lundblad and Michael Griffith for useful conversations pertaining to the Alabama variant, Dr Claudia Noyes for sharing her amino acid sequence data, and Dr Walter Kisiel for performing the analysis for β-hydroxyaspartic acid. We are especially grateful to Portia Hunt for typing the manuscript.

REFERENCES

17. Kisiel W: Personal communication
27. Noyes CM: Personal communication
Factor IXAlabama: a point mutation in a clotting protein results in hemophilia B

LM Davis, RA McGraw, JL Ware, HR Roberts and DW Stafford