**THE MOLECULAR BASIS OF HEMOPHILIA B:
IDENTIFICATION OF THE DEFECT IN FACTOR IX_{VANCOUVER}**

By

VALERIE ANNE GEDDES

B.Sc. (Hons) Simon Fraser University, 1984

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF**

MASTER OF SCIENCE

in

**THE FACULTY OF GRADUATE STUDIES
Department of Biochemistry
(Genetics Program)**

**We accept this thesis as conforming
to the required standard**

THE UNIVERSITY OF BRITISH COLUMBIA

February 1987

© Valerie Anne Geddes, 1987

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Biochemistry

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date 87:03:20

ABSTRACT

Hemophilia B_{Vancouver} is a moderately severe hereditary disorder in which the factor IX antigen is present in relatively normal amounts but the biological activity of factor IX is markedly reduced. Previous studies have demonstrated that although the patient has 62% of the normal factor IX antigen level in his plasma, he shows only 2.6% of normal factor IX procoagulant activity. In addition, radioimmunoassays have shown that epitopes on both the heavy and light chains of activated factor IX are present. These two results were taken as an indication that the molecular defect causing the hemophilia may be a point mutation involving an amino acid change in the protein.

In order to identify the mutation involved, DNA was isolated from lymphocytes in a blood sample from the patient. This DNA was used to construct a genomic library in the λ vector EMBL3. One million of the resultant clones were screened with a labeled factor IX cDNA probe to identify those clones containing portions of the factor IX gene. DNA inserts from three λ clones, which together span the entire gene, were subcloned to facilitate sequence analysis of the exons and intron / exon junctions of the factor IX gene. The nucleotide sequences of the coding regions were found to match the published sequence of the normal gene, except for one nucleotide. A single mutation was found at nucleotide 31,311 of the factor IX gene (Yoshitake et al., 1985), corresponding to amino acid 397 of the mature protein. This alteration, which changes an isoleucine codon, ATA, to a threonine codon, ACA, is novel among the mutations which have been reported to cause hemophilia B. A three dimensional model of the protease domain of factor IX_a, which was prepared on the basis of its homology to the pancreatic serine proteases, was examined in the vicinity of residue 397. The position of threonine 397 in this model suggests that this mutation could alter the hydrogen bonding between factor IX_a and its substrate, factor X. Taken together, these data suggest that this mutation is the cause of the hemophilia in this patient.

TABLE OF CONTENTS

ABSTRACT	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS	vii
ACKNOWLEDGEMENTS.....	ix
INTRODUCTION	p. 1
1) The Role of Factor IX in Blood Coagulation.....	p. 1
2) The Structure of the Factor IX Gene	p. 2
3) The Structure of the Factor IX Protein	p. 4
4) Hemophilia B as a Heterogeneous Disorder.....	p. 8
5) Classifications of Hemophilia B	p. 9
6) Mutations Known to Cause Hemophilia B	p. 10
7) Profile of the Subject's Family:	
A. Pedigree	p. 13
B. Factor IX Activity and Antigen Assays.....	p. 13
C. Classification of the Subject's Hemophilia	p. 16
8) Objectives of This Study	p. 17
MATERIALS & METHODS	p. 19
1) Materials	p. 19
2) Strains, Vectors, and Media	p. 20
A. Bacterial Strains	p. 20
B. Vectors.....	p. 20
C. Media	p. 20
3) Basic Molecular Biological Techniques	p. 21
A. Gel Electrophoresis	p. 21
B. DNA Isolations	p. 22
C. Production of DNA Fragments for Ligations.....	p. 25
D. Ligation of DNA into M13 Vectors	p. 26
E. Transformation of DNA into Bacteria.....	p. 27
F. Klenow Labeling of DNA Fragments.....	p. 28
4) Construction of the Genomic Library	p. 28
A. Isolation of DNA from Lymphocyte Nuclei.....	p. 28
B. Partial Digestions of Genomic DNA with Sau3A	p. 30
C. Ligations into λ EMBL3 Arms	p. 31
D. Packaging of Reconstituted EMBL3	p. 32
5) Screening of the Genomic Library.....	p. 32
A. Plating the Phage Library.....	p. 32
B. Hybridizations and Washing.....	p. 33

MATERIALS AND METHODS (cont'd)

6) Mapping of the Factor IX EMBL3 Clones.....	p.34
7) Subcloning into M13 vectors for Sequencing.....	p.35
A. Construction of M13 Clones.....	p.35
B. Screening of M13 Clones	p.37
C. M13 DNA Isolation.....	p.37
8) DNA Sequencing	p.38
9) Genomic Southern Blot.....	p.40

RESULTS

	p.42
--	------

1) Southern Blot Analysis of the SB IX-4 Factor IX Gene.....	p.42
2) Genomic Factor IX Clones Isolated from the SB IX-4 Library.....	p.42
3) M13 Subclones Containing SB IX-4 Factor IX Exons	p.44
4) Nucleotide Sequence of SB IX-4 Activation Peptide and Intron / Exon Junctions	p.44
5) Nucleotide Sequence Alteration in SB IX-4 Factor IX Gene.....	p.49

DISCUSSION.....

	p.52
--	------

1) Summary of Data Concerning Patient SB IX-4.....	p.52
2) Naming of the Mutation Found in SB IX-4	p.52
3) Implications for Carrier Detection and Prenatal Diagnosis.....	p.52
4) Possible Effect of the Ile ₃₉₇ to Thr ₃₉₇ Substitution on Factor IX _a Activity:	
A. Structure of Normal Factor IX _a	p.54
B. Structural Homology Between the Trypsin-Like Serine Proteases....	p.54
C. Model of the Complex Between Factor IX _a and Its Substrate Peptide	p.58
D. Hydrogen Bonding in the Normal Factor IX _a - Substrate Complex ..	p.62
E. Hydrogen Bonding in the Factor IX _a Vancouver - Substrate Complex .	p.62
5) Directions for Further Research.....	p.63
6) Summary.....	p.65

REFERENCES.....

	p.66
--	------

LIST OF TABLES

TABLE I.	Point mutations known to cause hemophilia B.....	p.11
TABLE II.	Factor IX clotting activity and antigen assays of the subject's family	p.15
TABLE III.	Composition of dideoxy / deoxy sequencing mixes.....	p.39
TABLE IV.	Nucleotide sequences of the SB IX-4 activation peptide and intron / exon junctions	p.48
TABLE V.	Amino acid sequence alignment of normal human factor IX _a and the pancreatic serine proteases.....	p.55

LIST OF FIGURES

FIGURE 1. Structure of the human factor IX gene, mRNA and protein.....	p.3
FIGURE 2. Pedigree of the subject's family with respect to hemophilia B	p.14
FIGURE 3. Southern blot analysis of the SB IX-4 factor IX gene	p.43
FIGURE 4. Genomic factor IX clones isolated from the SB IX-4 phage library ...	p.45
FIGURE 5. M13 subclones containing SB IX-4 factor IX exons	p.46
FIGURE 6. Nucleotide sequence alteration in the SB IX-4 factor IX gene.....	p.50
FIGURE 7. Hydrogen bonding in the factor IX _a - substrate complex models	p.60

LIST OF ABBREVIATIONS

A	Adenosine
ATP	Adenosine triphosphate
bp	Base pair(s)
BSA	Bovine serum albumin
C	Cytidine
dH ₂ O	Distilled water
DMSO	Dimethylsulfoxide
dNTP	Deoxyribonucleoside triphosphate
ddNTP	Dideoxyribonucleoside triphosphate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
G	Guanosine
Gla	γ - carboxyglutamic acid
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethane sulfonic acid
IPTG	Isopropyl- β -D-Thiogalactopyranoside
kb	Kilobase pairs
rpm	Revolutions per minute
mA	Milliamperes
mRNA	Messenger RNA
N	Nucleoside (G, A, T, or C)
OD	Optical density
PEG	Polyethylene glycol

pfu	Plaque forming unit
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodium dodecyl sulfate
TEMED	N, N, N', N'-Tetramethylethylenediamine
Tris	Tri (hydroxymethyl) aminomethane
tRNA	Transfer RNA
U	Units of enzyme activity
UV	Ultra violet
V	Volts
T	Thymidine
W	Watts
Xgal	5-Bromo-4-chloro-3-indolyl- β -D-galactopyranoside

ACKNOWLEDGEMENTS

I offer special thanks to the following friends and colleagues for their gifts of advice, reagents, encouragement and humor:

Gordon Louie

Bruce Tiberiis

Debbie Cool

Colin Hay

Dave Irwin

Kathy Robertson

Gary Brayer

I am indebted to Ross MacGillivray for his colossal generosity.

My most earnest thanks to my extraordinary mate, Chris Dewhurst.

INTRODUCTION

The arrest of bleeding following trauma is controlled by three interrelated factors: the constriction of blood vessels at the site of injury, the formation of a platelet plug, and the coagulation of blood. The process of coagulation occurs as a series of complex steps which terminate in the formation of a fibrin clot. Coagulation involves at least 14 plasma proteins, Ca^{2+} , phospholipid membrane surfaces, at least one tissue protein, and platelets (Jackson and Nemerson, 1980). Two systems exist for the activation of blood coagulation: the intrinsic pathway, which is a relatively slow process, and the extrinsic pathway, which is a much faster process (Hirsch and Brain, 1979). The protein coagulation factors, which circulate in the form of inactive precursors, are converted to their active forms in a cascade of sequential reactions.

The coagulation disorder known as hemophilia B is inherited as an X-linked recessive trait occurring approximately once in every 25,000 male births (Thompson, 1986). The disease is also known as Christmas disease and as plasma thromboplastin component deficiency. The clinical symptoms of hemorrhage in hemophilia B patients are due to the absence or decreased activity of blood coagulation factor IX.

1. The Role of Factor IX in Blood Coagulation

Factor IX circulates in plasma as a 54 kDa glycoprotein (Thompson, 1986), which constitutes 1/10,000 of the total circulating protein (Hedner and Davie, 1982).

During the middle phase of blood coagulation, factor IX is converted from its zymogen form to an active serine protease called factor IX_a . This activation can be achieved in two ways. As part of the intrinsic pathway of blood coagulation, factor IX is activated in a reaction with factor XI_a in the presence of calcium. This is the earliest Ca^{2+} dependent reaction in the intrinsic clotting system. An alternate route to

activation has been identified and involves the extrinsic coagulation system (Osterud and Rapaport, 1977). In this reaction, factor VII_a, in the presence of the lipoprotein known as tissue factor, activates factor IX. The relative importance of these two activation mechanisms is not understood (Thompson, 1986). Both of these reactions have the same biochemical result: the removal of an activation peptide leaves two polypeptide chains held together by disulfide bond(s). Activated factor IX acts as a serine protease to cleave factor X in the presence of Ca²⁺, phospholipid and factor VIII_a. The chain of activation reactions then continues until finally a fibrin clot is formed.

2. The Structure of the Factor IX Gene

The genomic and protein structures of factor IX are presented schematically in Figure 1. The gene for factor IX spans approximately 34 kilobases of DNA. Cytogenetic studies have localized the gene to the subtelomeric region of the long arm of the X chromosome. The factor IX gene is proximal to the gene for blood coagulation factor VIII and these two genes flank the site associated with fragile X mental retardation, at Xq27.3 (Purrello *et al.*, 1985; Camerino *et al.*, 1984; Chance *et al.*, 1983; and Boyd *et al.*, 1984).

The cloning and sequencing of the entire human factor IX gene (Yoshitake *et al.*, 1985; earlier work by Anson *et al.*, 1984; Kurachi and Davie, 1982; and Choo *et al.*, 1982) have shown that the mRNA is encoded in eight separate exons. The gene codes for a 2802 nucleotide mRNA (Anson *et al.*, 1984), of which about one half is a long 3' untranslated tail. Analysis of the genetic control elements by Yoshitake *et al.* (1985) resulted in the following observations. Putative promoter sites 5' to the factor IX coding region include residues -27 to -24. The sequence TGTA, which may

Figure 1
Structure of the human factor IX gene, mRNA and protein

The genomic structure of human factor IX (Yoshitake et al., 1985; Thompson, 1986) is represented on the top line from 5' (left) to 3' (right). Numbers below the line refer to kilobases. Exons 1 through 8 are indicated above the gene. Introns are indicated by capital letters. The mRNA is presented in the central diagram. The lighter section of the mRNA signifies the translated region and the darker section the long 3' untranslated tail. The translation product of this mRNA is shown in the lower diagram. Amino acids of the mature protein are numbered below the line. The contributions from each of the eight exons are indicated above the protein. Abbreviations used: "prepro" refers to the 46 amino acids between the translation initiation methionine and the processing cleavage site; "gla" refers to the domain containing gamma-carboxyglutamic acid residues; "EGF" stands for the epidermal growth factor homologous regions; "AP" for the activation peptide which contains asparagine-linked carbohydrate moieties and the two factor IX activation cleavage sites (indicated by arrows); and "tryptic" for the catalytic portion, including the active site serine, "S".

represent a TATA box (Goldberg-Hogness box; Breathnach and Chambon, 1981), is found at this site. This occurs 27 bp upstream of the initiation site as determined by nuclease S1 mapping analysis (Anson *et al.*, 1984). Two other sites further upstream of this region also show TATA homology similar to that observed in other genes, such as factor VIII. It has not been established whether these sites are associated with RNA polymerase II binding and the start of transcription. None of these sequences displays a nearby CCAAT regulator sequence (Breathnach and Chambon, 1981) nor upstream G/C rich regions (McKnight and Kingsbury, 1982).

The intron / exon junctions in the factor IX gene have been shown to conform to the GT / AG rule of donor and acceptor splice sites (Breathnach and Chambon, 1981) and to be very similar to the consensus sequence summarized by Mount (1982). A recognition sequence for poly(A) addition is found at position 32,741. It is identical to the AATAAA consensus sequence (Proudfoot and Brownlee, 1976) found to be necessary but not sufficient for poly(A) primary site selection. In the factor IX gene, the AATAAA sequence is found 39 bp upstream of the sequence CATTG. This sequence is homologous to the consensus sequence CA(T/C)TG, which may provide further specificity for cleavage site selection (Berget, 1984; Birnstiel *et al.*, 1985).

3. The Structure of the Factor IX Protein

The factor IX mRNA codes for a protein containing a prepro leader sequence which is not found in the mature plasma protein. Both the pre and pro sequences are removed during the biosynthesis of factor IX in hepatocytes. The hydrophobic pre sequence (residues -46 to -19) has limited homology to the signal sequences observed in other vitamin K - dependent clotting factors (Fung *et al.*, 1985). It is thought to be removed by a signal peptidase during translocation of the nascent protein across the

rough endoplasmic reticulum (von Henji, 1983 and 1985). The pro region (residues -18 to -1) is thought to be removed by an unidentified processing peptidase (Steiner *et al.*, 1980), after further modification of the protein.

The amino terminal domain of the mature plasma protein comprises the gla region, which takes its name from the γ - carboxyglutamic acid residues it contains. A vitamin K - dependent carboxylase system in the liver converts 12 glutamic acid residues to γ - carboxyglutamic acid. Factor VII, factor X, prothrombin, protein C, protein S and protein Z also contain gla domains (Suttie, 1985). The gla residues are associated with the sixteen Ca^{2+} binding sites observed in human factor IX (Bajaj, 1982), although one of these sixteen sites has been shown to be gla-independent (Morita *et al.*, 1984). The binding of Ca^{2+} to the strongly acidic γ - carboxyglutamic acid residues is a cooperative phenomenon which results in a conformational change in the molecule (Bajaj, 1982). The association of Ca^{2+} with factor IX greatly increases the rate of activation of factor IX by factor XI_a . Factor IX_a must interact with a phospholipid membrane, as well as with factor VIII_a and Ca^{2+} , in order to activate factor X. The γ - carboxyglutamic acid residues are thought to allow factor IX and other blood coagulation factors to form Ca^{2+} bridges to phospholipid membranes (Suttie, 1985).

The region of the protein directly following the gla residues contains two copies of a region showing homology to the murine epidermal growth factor (EGF) protein (Doolittle *et al.*, 1984; Blomquist *et al.*, 1984). These two factor IX regions are similar to 10 internally homologous subregions of the EGF protein. The same type of domain is found in several of the blood coagulation factors, including factors VII, X, XI and XII. It has been postulated that this region might have a role in the

interaction of factor IX with cell surfaces, since epidermal growth factor has this property *in vivo* (Anson *et al.*, 1984), but no evidence of this has been reported. The unusual amino acid β - hydroxyaspartate is also found in the EGF region. Its position at residue 64 is homologous to its position in protein C, factor X, and factor VII (McMullen *et al.*, 1983). The significance of this modification of an aspartate residue is unknown. It is possible that the β - hydroxyaspartate may be associated with the high affinity gla-independent Ca^{2+} binding site (Morita *et al.*, 1984).

The sites for the activation cleavages of factor IX are found in the region of the protein downstream of the EGF domain. The bonds cleaved are those between arginine₁₄₅ and alanine₁₄₆ and between arginine₁₈₀ and valine₁₈₁. After these cleavages occur, the amino terminal chain and the carboxy terminal chain are held together by a disulfide bond or bonds. The activation peptide removed during this process is 35 amino acids in length and contains two of the asparagine-linked carbohydrate moieties found in factor IX. During the activation of factor IX by factor XI_a, the arginine - alanine bond is cleaved first. The product of this cleavage, factor IX _{α} , has no proteolytic activity. Some *in vitro* activators of factor IX, such as Russell's viper venom, cleave the arginine-valine bond at a greater rate than the arginine-alanine bond. The product of the cleavage of the arginine-valine bond only, factor IX_{ac}, has some activity (Lindquist *et al.*, 1978). The fully active species generated when both of the bonds have been cleaved is referred to as factor IX_{a β} . By analogy with the pancreatic serine proteases, the free amino group generated by the activation reaction (valine 181) is likely to be involved in an ion pair interaction with the free carboxyl group of the aspartate residue at position 364 (Thompson, 1986). This association probably induces a conformational change in the catalytic

subunit of factor IX_{aβ}. The change in conformation may consolidate the molecule into a configuration which increases the reactivity of the active site serine (Fujikawa *et al.*, 1974; Furie *et al.*, 1982).

The serine protease domain is found in the heavy chain of factor IX_a. The catalytic subunit includes the active site serine (position 195 in chymotrypsin), aspartic acid (102) and histidine (57) residues which are thought to contribute to the charge relay system (Kraut, 1977). There is a close evolutionary relationship between the factor IX_a catalytic subunit and members of the trypsin-like serine protease family. This group includes the pancreatic serine proteases chymotrypsin, trypsin, and elastase as well as kallikrein, plasmin, thrombin, and blood clotting factors XII_a, X_a, VII_a and XI_a (Jackson and Nemerson, 1980).

In addition to having the ability to cleave a specific peptide bond, there are a number of other requirements for the full biological activity of factor IX. The pre and pro leader sequences must be removed and the protein must be secreted into the bloodstream. Modifications of the gla residues and the β - hydroxyaspartate residue must occur during maturation of the protein in hepatocytes. The gla region must effectively bind Ca²⁺ and must be able to associate with phospholipid membranes. Factor IX must be recognized by factor XI_a. Both activation cleavages must occur to obtain a fully active protease. Finally, factor IX_a must be able to interact in some manner with its protein cofactor, factor VIII_a.

4. Hemophilia B as a Heterogeneous Disorder

Since factor IX has a complex role in coagulation, it is not surprising that the profiles of patients with hemophilia B vary widely with respect to clinical severity, clotting factor activities and antigen levels (Thompson, 1986). The wide heterogeneity of clinical symptoms was first recognized by Fantl (1956) and Roberts *et al.* (1968). These authors suggested that the disease existed in at least two forms. Roberts found that for 10% of his patients, hemophiliac plasma could neutralize factor IX inhibitors. This result suggests that their plasma contained significant amounts of factor IX antigen. The remaining 90% of his patients did not display such neutralization of inhibitors.

Further investigations (Hougie and Twomey, 1967; Gray *et al.*, 1968) identified a group of patients exhibiting a prolonged prothrombin time when ox brain, rather than human brain, was used as a source of thromboplastin reagent. Such hemophilias were referred to as hemophilia B_m. The prothrombin time assay (abbreviated PT) is a standard hematological measure of the activity of the extrinsic pathway. The majority of hemophilia B patients give a normal prothrombin time with both human and ox brain thromboplastin. Brown *et al.* (1970) noted that hemophiliacs giving a prolonged ox brain PT were severely affected and clinically indistinguishable from one another. Those patients not exhibiting the prolonged ox brain PT time were distinguishable in terms of clinical profile.

A more extensive survey of hemophilia B patients by Kasper *et al.* (1977) showed that hemophilia B is even more heterogeneous a disease than was previously recognized. Of 92 patients surveyed, 36 showed a discrepancy between antigen and activity levels, while 56 showed a proportional decrease in these two parameters. The

authors divided their patients into six groups with respect to their clinical severity, factor IX activity, antigen levels, and ox brain PT results. Within several of these groups, however, there was further evidence of heterogeneity. All hemophilia B_m patients in this study had an excess of antigen over activity. However, there was evidence of heterogeneity within the B_m group with respect to the degree of prolongation of the ox brain PT time. Another study has shown that one variant of factor IX_{Bm} acts as a competitive inhibitor of factor X when the latter is activated by the reaction product of bovine tissue factor and human factor VII (Osterud *et al.*, 1981). A recent paper reported three variants of factor IX_{Bm} which all exhibit different rates of activation cleavage (Yoshioka *et al.*, 1986). Thus, the B_m phenotype does not appear to define a specific subgroup within the disease.

5. Classifications of Hemophilia B

Studies such as the above have resulted in the development of several classification systems for hemophilia B variants. Hemophilias are identified as mild (over 4% of normal factor IX clotting activity), moderate (2 - 4%) or severe (less than 2%). Hemophilias also vary with respect to the presence or absence of circulating inhibitors of factor IX. These inhibitors are thought to be alloantibodies raised during treatment with exogenous factor IX. They are presumed to arise when the patient's own factor IX is either present at a very low level, contains deletions, or is heavily modified (Ratnoff, 1978). Another classification of hemophilia B variants is based on the presence or absence of factor IX antigen (Roberts *et al.*, 1968). Those patients displaying normal levels of antigen but reduced levels of factor IX activity are designated crm⁺ (cross reacting material positive); those displaying reductions in both activity and antigen, crm^{red}; and those showing undetectable antigen and severely reduced activity, crm⁻. Crm⁺ mutations and those crm^{red} hemophilias with a large

discrepancy between antigen and activity levels are of particular interest, since these mutations must result in critical changes to the factor IX protein which affect its biological function.

6. Mutations Known to Cause Hemophilia B

The heterogeneity in the clinical profiles of hemophilia B is a strong indication that the disease is caused by a variety of genetic mutations. Knowledge of the sequence of the factor IX gene and of the sequence and characteristics of the protein has led to investigations into the relationship between structure and function in factor IX. Table 1 presents the point mutations causing hemophilia B which have been identified at the amino acid or nucleotide level.

Two mutations have been identified which involve alterations in the prepro leader sequence. Factor IX_{Oxford}, associated with a severe form of hemophilia B, exhibits an amino acid change at residue -4 from an arginine residue to a glutamine (Bentley *et al.*, 1986). This alteration appears to preclude the removal of the 18 amino acid propeptide and results in decreased factor IX activity in the patient's plasma. Factor IX_{Oxford} was found to bind calcium normally and to be γ carboxylated to a normal degree. The authors suggest that either the basic charge on the propeptide or steric interference due to the extra 18 amino acids could alter phospholipid binding or other properties.

Factor IX_{Cambridge} contains an alteration from an arginine at position -1 to a serine residue (Diuguid *et al.*, 1986). Like Factor IX_{Oxford}, this mutation results in a factor IX which is extended by the 18 amino acids of the propeptide. The failure to remove the leader peptide appears to have a profound effect on the activity of factor

Table I.
Point mutations known to cause hemophilia B

Mutation name	residue	mutation	region	Factor IX levels antigen	activity
1) Factor IX _{Oxford}	aa - 4	arg to gln	prepro	89%	<0.5%
2) Factor IX _{Cambridge}	aa - 1	arg to ser	prepro	80%	<1%
3) Factor IX _{Alabama}	aa 47	asp to gly	EGF	-	moderate
4) Factor IX _{Chapel Hill}	aa 145	arg to his	AP	normal	8%
5) donor splice site	n 21,165	GT to TT	exon 6	0.3%	<0.5%

The location of each mutation is indicated in the second column, followed by the amino acid and/or nucleotide change observed (aa = amino acid, n = nucleotide). The region of the factor IX protein in which the mutation occurs is also indicated. Clinical symptoms, as far as they were reported in the original papers, are also recorded. Antigen levels were measured by immunoassay and activity levels by standard one or two stage clotting time assays. References: 1) Bentley et al., 1986; 2) Diuguid et al., 1986; 3) Davis et al., 1984; 4) Noyes et al., 1983; and 5) Rees et al., 1985.

IX_{Cambridge}. The presence of the abnormal amino terminal extension appears to affect the carboxylation of the gla residues, suggesting that the sequence prior to the gla domain is important in recognition by the vitamin K - dependent carboxylase system (Bentley *et al.*, 1986).

The molecular defect in factor IX_{Alabama} has been reported as an aspartate to histidine change in the EGF region of the protein. When the sequence of the protein was analyzed (Davis *et al.*, 1984), no other change from the normal sequence was observed. The amino acid change results in the loss of a negative charge. As the role of the EGF domain itself is not understood, the role of this residue in the protein is not clear.

In the hemophilia associated with factor IX_{Chapel Hill}, amino acid 145 has been changed from an arginine residue to a histidine residue. Arginine₁₄₅ is the site of one of the activation cleavages of factor IX. Both factor XI_a and VII_a require the arginine side chain in this position to recognize the cleavage site. Thus, the change to a histidine precludes activation. Since the partially activated factor IX_{aa} has reduced biological activity, this mutation lowers factor IX_a activity to 8% of normal.

A splicing mutation is responsible for the decreased activity of factor IX in another case of hemophilia B (Rees *et al.*, 1985). DNA analysis has shown that in this variant, the donor splice site GT of exon 6 has been altered to a TT dinucleotide. Such a mutation would be expected to result in an altered mRNA. Several different types of RNA can result from splicing mutations: either the entire intron is retained, or cryptic splice sites are used. Often splicing mutations result in truncated messages or in mRNAs which are unstable. Such mutations have previously been observed in

the β^0 thalassemias (Weatherall *et al.*, 1984). The precise effect of the splicing mutation in this hemophilia has not been determined. However, it is known that the mutation results in an almost undetectable factor IX antigen level in the plasma.

Many more variants of hemophilia B have been identified which involve partial or gross deletions of the factor IX gene. Some of these studies have identified the boundaries of these deletions, while others have identified them only in a qualitative way (Bray and Thompson, 1985, 1986; Hassan *et al.*, 1985; Vidaud *et al.*, 1986; Giannelli *et al.*, 1983). In addition to these mutations, there is a further group of patients for which a molecular diagnosis has not been obtained (Thompson, 1983; Mertens *et al.*, 1983; Bertina and van der Linden, 1982a; Usharani *et al.*, 1985; Briet *et al.*, 1982; Osterud *et al.*, 1979). For these patients, analyses of blood clotting characteristics and of the presence or absence of inhibitors of factor IX are often the only data available. These patients present a wide variety of clinical profiles.

7. Profile of the Subject's Family:

A. Pedigree

A British Columbian family with a history of hemophilia B has undergone carrier detection using restriction fragment length polymorphisms (Hay *et al.*, 1986). The pedigree of this family with respect to hemophilia is found in Figure 2. Hemophilia B has been observed over several generations, indicating that the mutation in patient SB IX-4 is more likely to have been inherited than to be a new mutation.

B. Factor IX Activity and Antigen Assays

During the carrier detection study, factor IX clotting activity and factor IX antigen levels were measured for several members of the family. Table II lists the results of these assays. In each case, the values are expressed as a percentage of the 'normal'

Figure 2
Pedigree of the subject's family with respect to hemophilia B

Affected males are indicated by filled squares and suspected carrier females by dotted circles. Normal individuals are indicated by white circles (females) or white squares (males). Question marks indicate individuals for whom carrier status has not been established. The hemophiliac whose gene is under study has been indicated by an arrow and identified as individual SB IX-4.

Table II

Factor IX clotting activity and antigen assays of the subject's family

Subject	IX:C	IX:Ag a	IX:Ag b
Father	110	98	113
Mother (carrier)	18	73	70
Son (hemophiliac), SB IX-4	2.6	61	63
Daughter	100	92	90
Control	108	104	121

The factor IX clotting activities (IX:C) of members of the hemophilia B family and an unrelated, unaffected male were measured with a one stage assay using factor IX deficient plasma. The factor IX antigen levels (IX:Ag) were assayed using two methods: a, using a polyclonal non-calcium-dependent preparation for both the coating and the labeled species; and b, using a monoclonal antibody to an epitope on the heavy chain of factor IXa (Thompson, 1983) as the solid phase and a calcium-dependent polyclonal fraction which recognizes a light chain epitope (Bray et al., 1986) as the labeled species. Results are expressed as percent of normal values. From Hay et al., 1986.

value. 'Normal' refers to the mean value from many unaffected individuals. The immunoassays show that the hemophiliac son's circulating factor IX contains epitopes from both the light and heavy chains of factor IX_a (column IX:Ag b). In this assay, a monoclonal antibody to an epitope on the heavy chain of factor IX_a (Thompson, 1983) was used as the solid phase attached to microtitre wells. Added factor IX_a would therefore not be retained in the wells unless it contained the heavy chain epitope. A calcium - dependent polyclonal antibody fraction, which recognizes a light chain epitope (Bray *et al.*, 1986), was used as the labeled species. When this labeled preparation was applied, radioactivity would not be associated with the wells unless the retained factor IX_a presented the light chain epitope. While the patient's factor IX clotting activity was measured to be 2.6% of the normal level, his antigen level was shown to be considerably higher, at approximately 62% of normal.

C. Classification of the Subject's Hemophilia

The above assay results place the patient SB IX-4 in the moderate, crm^{red} class of hemophilia B variants. No ox brain prothrombin test has been performed on the patient's plasma to date. Plasma from SB IX-4 does not contain circulating inhibitors to factor IX (G. Growe, personal communication). The assay results suggest that the molecular defect responsible for hemophilia B in this family is likely to be a small deletion or point mutation within the protein coding region of the factor IX gene. Changes such as major gene deletions, splicing mutations, and mutations in transcriptional control elements would be expected to reduce antigen and activity levels proportionately.

The reason for the reduction in the factor IX antigen level in SB IX-4 to 62% of normal was not under investigation in this study. Unequal reduction in activity and antigen level is a common feature in patients with hemophilia B (Kasper *et al.*, 1977;

25 patients of 92 investigated). In addition, there is a very wide range of factor IX activity values even among normal individuals: a range of 37 to 170% was noted in one survey (Ratnoff, 1978). Within the family under study, however, clotting activities correlate with carrier or hemophilic status. Hemophilias with unequal reduction in activity and antigen are not well understood. The suggestion has been that an amino acid change in the protein could either alter the clearance of circulating factor IX from the blood or reduce the ability of the antibodies used in the immunoassay to recognize factor IX_a. Two distinct radioimmunoassays, which used antibody preparations recognizing different epitopes on the factor IX molecule, were performed. Since these assays gave similar results, it seems unlikely that epitope changes could explain the reduced antigen value. The reduction in factor IX antigen level is probably due to an altered half-life for the mutant protein.

8. Objectives of This Study

The purpose of the present study was to determine the molecular defect resulting in decreased specific coagulation activity of the factor IX found in SB IX-4. An analysis of such a defect can be performed in several ways. One method to pinpoint a mutation is to prepare a cDNA library and sequence the factor IX cDNA. This approach, however, is impractical in a human study. Since factor IX is synthesized in hepatocytes, a liver sample would be required to provide an RNA source containing factor IX mRNA. Another approach is to isolate and sequence the mutant protein. The third approach, and the one chosen for this study, is to prepare a genomic library from the patient's DNA. DNA isolated from peripheral lymphocytes in a blood sample from SB IX-4 was used to prepare a genomic library. The library was screened by using a normal human factor IX cDNA as a hybridization probe.

The DNA inserts from positive clones were then subcloned and the exons and intron / exon boundaries of the gene were sequenced. In this manner, a coding region mutation in the serine protease subunit of factor IX_a was identified.

MATERIALS and METHODS

1. Materials

Agarose, acrylamide, bisacrylamide, urea, ammonium sulfate and TEMED were obtained from Bio-Rad Laboratories. Yeast extract, casamino acids, bacto-tryptone, and bacto-agar were Difco grade and obtained from the Grand Island Biological Company. Nitrocellulose sheets and circles (82 and 132 mm, 0.45 μ m pore size) were obtained from Millipore or Schleicher & Schuell. 32 P - labeled nucleotides were obtained from Amersham. Phenol was obtained from the British Drug Houses Ltd and was redistilled before use. The fraction distilled at 179° C was collected and frozen in aliquots at - 20 °C. Deoxy- and dideoxy-ribonucleotides were purchased from PL-Pharmacia. IPTG, Xgal, EtBr, DMSO, yeast tRNA, chloramphenicol and ribonuclease A were obtained from Sigma. Cesium chloride was purchased from Cabot Berylco Ltd. Ultrogel AcA54 was purchased from LKB. All other chemicals were of reagent grade or higher and were purchased from either Sigma Chemical Co., Fisher Scientific, or British Drug Houses Limited. Restriction endonucleases, T4 DNA ligase, T4 DNA polymerase and nuclease-free BSA were obtained from New England Biolabs, Bethesda Research Laboratories, or PL-Pharmacia. DNA Polymerase I Klenow fragment was obtained from PL-Pharmacia. Blood samples were taken by Dr. G. H. Growe, Department of Medicine, Vancouver General Hospital.

2. Strains, Vectors and Media

A. Bacterial Strains

E. coli LE 392 and its P2 lysogen P2 392 (F⁻, hsd R514, (r_K⁻, m_K⁺), sup E44, sup F58, (lacY1 or Δ (lac IZY)), galK2, galT22, metB1, trpR55, λ⁻; Maniatis *et al.*, 1982) were hosts for the screening and isolation of clones in EMBL3. *E. coli* JM103 (Δlacpro, strA, thi⁻, supE, sbcB15, endA, hsdR⁻, F', traD36, proAB, lacI^q, lacZΔM15; Messing, 1983) was the host for transformation and isolation of clones in M13. *E. coli* MC1061 was used as the host for the factor IX cDNA clone (pCHIX).

B. Vectors

The genomic library was constructed in the phage λ vector EMBL3 (Frischauf *et al.*, 1983). For DNA sequence analysis, the M13 vectors mp10, 18 and 19 were used (Messing, 1983). The factor IX cDNA (pCHIX) was previously cloned into the plasmid vector pKT218.

C. Media

The medium for growth and screening of λ clones and hosts was NZYM (Maniatis *et al.*, 1982; 10 g NZ Amine type A, 2 g MgCl₂, 5 g NaCl and 5 g yeast extract per liter, pH adjusted to 7.5 with NaOH). For screening λ libraries, the phage were plated on NZYM agar (1.5%, w/v) plates with an overlay of NZYM agarose (0.7%, w/v). For titering of phage stocks, the overlay consisted of NZYM agar in place of NZYM agarose.

Bacteria containing M13 clones were grown in YT medium (5 g yeast extract, 8 g bacto-tryptone, and 5 g NaCl per liter, pH 7.5). Phage M13 transformants were plated on YT agar (1.5%, w/v) plates overlaid with YT agarose (0.7%, w/v). *E.*

coli JM103, the host for M13 vectors, was maintained on minimal medium plates (Messing, 1983). Minimal medium was prepared as follows: 3 g of agar in 160 ml dH₂O was autoclaved, cooled to 55 °C, and was mixed with 40 ml 5X salts (2.1 g K₂HPO₄, 0.9 g KH₂PO₄, 0.2 g (NH₄)₂SO₂, 0.1 g Na citrate·7H₂O per 40 ml). Autoclaved solutions of 20% glucose (2 ml), 20% MgSO₄ (0.2 ml), and a filter sterilized solution of 10 mg/ml thiamine (0.1 ml) were added. Plates were poured immediately.

3. Basic Molecular Biology Techniques

A. Gel Electrophoresis

DNA fragments were separated according to size by electrophoresis in agarose or polyacrylamide gels. The buffer for agarose gel electrophoresis was 1X TAE (50X TAE contains 2 M Tris base, pH 7.5, 1 M glacial acetic acid, and 0.1 M EDTA; Maniatis *et al.*, 1982). Agarose gel electrophoresis was performed in a horizontal apparatus. DNA fragments in these gels were visualized by fluorescence in the presence of EtBr (0.5 µg/ml) under UV irradiation at 260 nm. The buffer for polyacrylamide gel electrophoresis was 1X TBE (10X TBE contains 0.89 M Tris base, 0.89 M boric acid, 25 mM EDTA, pH 8.3; Maniatis *et al.*, 1982). Denaturing polyacrylamide gels contained 8.3 M urea; non-denaturing polyacrylamide gels did not contain urea. For non-denaturing gels, acrylamide (added to the appropriate concentration from a stock solution of acrylamide:bisacrylamide in dH₂O (38% : 2%, w/v) and buffer were mixed with the appropriate volume of dH₂O, and degassed.

Polymerization was initiated by the addition of ammonium persulfate and TEMED to final concentrations of 0.066% (w/v) and 0.04% (w/v), respectively. The gels were stained with 0.5 µg/ml EtBr in dH₂O for 10 min, and the DNA was visualized under

UV irradiation. Polyacrylamide gels in TBE buffer containing urea were prepared, degassed and polymerized in the same manner except that the final concentration of TEMED was 0.024% (w/v). Polyacrylamide gel electrophoresis was performed in a vertical apparatus without pre-electrophoresis. DNA in denaturing gels was visualized by autoradiography after drying the gels under vacuum in a Bio-Rad gel drier at 80 °C for 20 - 30 min, followed by exposure to Kodak XK-1 film, with or without intensifying screens.

B. DNA Isolations

Precipitations of DNA from aqueous solution was achieved by the addition of 1/10 volume of 3 M NaOAc and either 2 volumes of - 20 °C 95% ethanol or 1 volume of room temperature isopropanol. The resulting solution was mixed thoroughly and incubated at either 0 °C, - 20 °C, or - 70 °C unless otherwise specified. Incubation time at these temperatures varied from 30 min to overnight. After incubation, the solutions were centrifuged in either a microfuge at top speed or a Sorvall centrifuge at 10 krpm for 30 min. Nucleic acid pellets were washed once with - 20 °C 75% ethanol and allowed to air dry.

Double stranded replicative form M13 DNA was isolated as described by Messing (1983). A single plaque was used to inoculate 2 ml of YT containing 10 µl of an overnight JM103 culture. Small amounts of M13 replicative form DNA were prepared by a modification of the method of Birnboim and Doly (1979) (Maniatis *et al.*, 1982). An aliquot (1.5 ml) of the 6 h culture of the clone of interest was placed in a microfuge tube and the bacteria were collected by centrifugation for 1 min in a microfuge. The cell pellet was resuspended in 100 µl of an ice cold solution containing 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8.0, and 4 mg/ml

lysozyme. After the suspension was incubated for 5 min at room temperature, 200 μ l of a solution containing 0.2 N NaOH-1% SDS was added. The mixture was incubated for 5 min on ice, and 150 μ l of potassium acetate solution (60 ml 5 M KOAc, 11.5 ml glacial acetic acid, 28.5 ml dH₂O, pH 4.8) was added. After vortexing, the suspension was incubated at 4 °C for 5 min. Cellular debris was removed by centrifugation in a microfuge for 5 min at 4 °C. The supernatant was removed and extracted with an equal volume of phenol:chloroform (1:1, v/v). Nucleic acids were precipitated with ethanol at room temperature. The nucleic acids were resuspended in 50 μ l of TE buffer (10 mM Tris-HCl, pH 8.0, 1mM EDTA) containing 20 μ g/ml pancreatic RNase A .

A larger scale preparation of some M13 replicative form DNA was also performed. In this case, the 6 h culture grown from the single plaque was used to inoculate 100 ml of fresh YT. One milliliter of a culture of uninfected JM103 cells grown to log phase was added. The 100 ml culture was grown for a further 6 h. The previously described extraction procedure, scaled up by a factor of 6, was used to isolate the cloned DNA.

A large scale alkali lysis procedure was used to purify plasmid DNA. An aliquot from an overnight culture (5 ml) of pCHIX grown in *E. coli* MC1061 was used to inoculate 1 l of YT. The culture was incubated at 37 °C until the OD at 400 nm was > 0.4. At this time, 5 ml of chloramphenicol solution (34 mg/ml in ethanol) was added and the culture was incubated overnight at 37 °C. The extraction procedure is a scaled-up version of the previously described method with the following modifications. After the addition of potassium acetate, the debris was removed by centrifugation at 35 krpm in a Ti60 rotor for 30 min at 4 °C. Plasmid DNA was

separated from chromosomal DNA and RNA by isopycnic centrifugation using a cesium chloride gradient. The gradient solutions were made by the addition of 3.9 g CsCl and 0.3 ml EtBr (10 mg/ml) to 3.8 ml of the resuspended DNA. The gradient was then formed by centrifugation at 50 krpm using a Ti70.1 Beckman rotor for 20 h. The bright red band was removed from the tube with a syringe. This solution was then extracted 5 times with equal volumes of 2-butanol and dialyzed for 3 h against several changes of 1 l TE at 4 °C. The nucleic acids were then precipitated from ethanol at - 20 °C overnight and resuspended into 300 µl of TE.

For large scale preparations of EMBL3 clones (Maniatis *et al.*, 1982), 2×10^{10} host bacterial cells grown in YT with 0.2% maltose were collected by centrifugation and resuspended in 4 ml of SM buffer (5.8 g NaCl, 2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 50 ml 1 M Tris-HCl, pH 7.5, 5 ml 2% gelatin (w/v) per liter). The cloned EMBL3 phage (2×10^7 pfu) were added to the cells, and the phage were allowed to adsorb to the cells during an incubation at 37 °C for 20 min. This mixture was used to inoculate 400 ml of prewarmed NZ medium and the culture was incubated at 37 °C until lysis, approximately 4 h. Chloroform (4 ml) was added to lyse the remainder of the cells by incubation at 37 °C for 10 min. The solution was brought to 1 M NaCl by the addition of 80 ml of 5 M NaCl. Bacterial debris was removed by centrifugation at 5 krpm in a GSA rotor for 10 min. The supernatant was then digested with DNase I (800 µl of 1 mg/ml solution) and RNase A (80 µl of 10 mg/ml solution) for 30 min at 4 °C. Phage particles were precipitated by the addition of 80 g of polyethylene glycol 6000 (Carbowax 8000) followed by incubation at 4 °C overnight. Phage particles were collected by centrifugation at 5 krpm in a GSA rotor for 10 min at 4 °C. After removal of the PEG / NaCl solution, the phage particles were gently resuspended in 25 ml SM. An extraction with an equal volume of CHCl_3 was followed by

centrifugation at 8 krpm for 10 min. Phage were purified from the aqueous supernatant in CsCl gradients. Gradients were made by the addition of 18 g CsCl per 25 ml of phage solution. Centrifugation at 60 krpm was continued for 16 - 20 h at 20 °C in a Ti70.1 rotor. Phage were removed from the gradient after visualization as a blue band visible in white light. The solution of recovered phage was made 0.02% with respect to gelatin. CsCl was removed by dialysis of the solution against 10 mM Tris base, pH 8.0, 25 mM NaCl, and 10 mM MgCl₂ for 2 h at 4 °C. SDS was added to 0.05% (w/v), EDTA to 20 mM, and proteinase K to 50 µg/ml. The solution was incubated at 68 °C for 1 h. The DNA was purified by extraction with phenol:chloroform (1:1, v/v), followed by an extraction with chloroform alone. The solution was dialyzed overnight against TE at 4 °C. Phage DNA was precipitated from ethanol at 0 °C for 30 min. The DNA was resuspended in 500 µl TE.

Small scale preparations of EMBL3 clones were scaled down from the large preparation described above (Maniatis *et al.*, 1982). Eluted phage from a single plaque were adsorbed to 100 µl of host cells at 37 °C for 10 min and used to inoculate 20 ml of NZ medium. DNA isolation followed the above protocol until the CsCl gradient step. At this point, the phage suspension was digested with proteinase K by the addition of SDS to 1%, EDTA to 5 mM, and proteinase K to 0.15 mg/ml. DNA was precipitated from isopropanol at 0 °C for 1 h and resuspended in 100 µl of TE buffer.

C. Production of DNA Fragments for Ligation

In order to isolate specific fragments or sizes of fragments for ligations or other manipulations, restriction enzyme digests were electrophoresed on 0.7% low melting

point agarose gels. For the long genomic fragments (10 - 20 kb), the volume of the gel slice was estimated and 5 volumes of 20 mM Tris base, pH 8.0, 1 mM EDTA was added to the same tube. This mixture was heated in a 68 °C bath until the gel had completely melted. When the solution had returned to room temperature, it was extracted with equal volumes of phenol, then a 1:1 mixture of phenol:sevag (24:1 chloroform : iosamylalcohol), and finally with sevag alone. The DNA was then precipitated from ethanol at room temperature for 20 min and resuspended in an appropriate volume of dH₂O.

For shorter fragments, a modification of this procedure was used. In this case, 500 µl or less of the agarose slice was placed in an eppendorf tube and melted by incubation at 68 °C. When the temperature of the gel had returned to approximately 40 °C, an equal volume of phenol was added to the tube, and the mixture was vigorously vortexed. The tube was then placed in a dry ice - ethanol bath at - 70 °C for 15 min. After the tube was removed and the solution completely thawed, the tube was centrifuged for 15 min in a microfuge. The aqueous layer was removed into a fresh tube. Approximately 150 µl of TE was added to the phenol layer in the original tube. This tube was then vortexed and centrifuged in the microfuge as previously described. The two aqueous layers were then pooled, and extracted once with an equal volume of phenol and once with sevag. After the extractions, the aqueous layer was transferred to a fresh eppendorf tube. The DNA was precipitated from ethanol at room temperature for 20 min.

D. Ligation of DNA into M13 Vectors

DNA fragments were ligated into vector DNA in 10 - 15 µl of a buffer containing 66 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 5 mM DTT, and 0.4 - 1 mM ATP. For

ligations, 10 - 20 ng vector DNA was incubated with a 1 - 5 molar excess of insert DNA. T4 DNA ligase was added at 1 U for blunt ended ligations and 0.1 U for sticky ended ligations (Maniatis *et al.*, 1982). Ligations were allowed to proceed overnight at 15 °C.

E. Transformation of DNA into Bacteria

Host bacteria for M13 transformations were made competent by treatment with CaCl_2 (Dagert and Ehrlich, 1979; Messing, 1983). Fifty millilitres of YT were inoculated with a host colony from a minimal glucose plate and incubated at 37 °C with vigorous shaking until the OD_{600} of the culture was 0.4 to 0.6. Cells were collected by centrifugation at 2.5 krpm in an HB-4 rotor at 4 °C for 5 min, and were gently resuspended in 1/2 of the starting volume of ice cold 50 mM CaCl_2 . The cells were incubated on ice for 25 min and were again collected by centrifugation in an HB-4 rotor at 2.5 krpm for 5 min at 4 °C. The bacteria were gently resuspended in 1/10 of the starting volume in ice cold 50 mM CaCl_2 . The highest transformation efficiency was obtained when the cells were stored at 4 °C for 24 h before use (Dagert and Ehrlich, 1979). Aliquots (0.3 ml) of competent cells were transformed with 2 μl of ligated DNA (see previous section). Cells were incubated with DNA in 13 X 100 mm glass tubes at 4 °C for 40 min and then heat shocked at 42 °C for 2 min. M13 transformants were plated immediately with the addition of 10 μl 100 mM IPTG, 50 μl Xgal (10 mg/ml in dimethylformamide), 0.2 ml host cells in log phase, and 4 ml soft YT agarose, and spread on YT plates. The plates were then incubated at 37 °C overnight. Recombinants were identified as colorless plaques in the presence of Xgal and IPTG (Messing, 1983).

F. Klenow Labeling of DNA Fragments

DNA was labeled using the method of Feinberg and Vogelstein (1983). Typical reaction mixtures contained 100 - 200 ng DNA in a volume of 50 μ l. DNA in 32 μ l of dH₂O was denatured by boiling for 10 min, and was cooled to 37°C for 15 - 30 min. Labeling was set up in a final volume of 50 μ l of 50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 10 mM β mercaptoethanol, 20 μ M dCTP, 20 μ M dGTP, 20 μ M dTTP, 50 μ Ci α - ³²P dATP (3000 Ci/mmol), 200 mM HEPES, pH 6.6, 60 A₂₆₀/ml p(dN₆) (random primers), 0.4 mg/ml BSA and 0.1 U/ μ l *E. coli* DNA Polymerase I (Klenow fragment). Extension was allowed to occur overnight at 37 °C. The reaction was terminated by the addition of SDS to 0.7% and EDTA to 7.5 mM. The total volume was brought to 200 μ l with dH₂O. Labeled DNA was separated from unincorporated labeled nucleotides by chromatography on an Ultrogel AcA54 column by elution from the column with a solution of 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, and 0.25 mM EDTA. The specific activity of the probes labeled in this manner was between 0.5 - 4 X 10⁸ cpm/ μ g. Labeled DNA was denatured immediately before use by boiling for 10 min or by addition of 1/10 volume 1 N NaOH and incubation at 68°C for 10 min. If the DNA was denatured with alkali, 1/10 volume of 1.5 M NaH₂PO₄ was added prior to the hybridization reaction.

4. Construction of the Genomic Library

A. Isolation of DNA from Lymphocyte Nuclei

The protocol for isolating DNA from peripheral lymphocytes was derived from that of Bell *et al.* (1981). Frozen whole blood collected in tubes containing EDTA was frozen at -20 °C in 20 ml aliquots. The blood was thawed by standing at room temperature for 30 min. It was then diluted with 200 ml of sucrose buffer (0.32 M

sucrose, 10 mM Tris-HCl, pH 7.5, 5 mM MgCl_2) and transferred to 250 ml centrifuge bottles. The solution was mixed gently for 5 min at room temperature, then centrifuged at 4 krpm, 4 °C, in a Sorvall GSA rotor for 10 min. The supernatant was carefully removed by aspiration. Sucrose buffer (25 ml) was then added to the pellet and the white blood cells and red blood cell debris were resuspended using a sterile pasteur pipet. The centrifugation step was repeated, and again the supernatant fraction was removed by aspiration.

The pellet was resuspended in 9 ml of nuclei buffer (75 mM NaCl, 24 mM EDTA, pH 8.0). The solution was transferred to a sterile 50 ml tube, and 225 μl of 20% SDS solution was added. The mixture was gently inverted and left to stand for 15 min at room temperature. A small amount of proteinase K powder was added to the solution (10 mg). The mixture was then incubated overnight at 37 °C.

Each sample was extracted 3 times with equal volumes of phenol. The tube was rocked gently for 30 min at room temperature. The mixtures were then centrifuged at 4 krpm in a Sorvall HB-4 rotor for 10 min at room temperature. The aqueous supernates were removed to a fresh tube. Two further phenol extractions were performed in the same manner. Finally, a sevag extraction was performed.

One tenth volume (800 μl) of 4 M NH_4OAc was added and the solution was thoroughly mixed for 45 min. An equal volume of isopropanol was added and the tube gently inverted. The filamentous DNA precipitate was collected on the end of a micropipet. The pellet was washed twice with 95% ethanol and allowed to air dry. The DNA was resuspended in 200 μl TE (10 mM Tris base, pH 8.0, 1 mM EDTA), and was allowed to rehydrate for one week before use.

B. Partial Digestions of Genomic DNA with Sau3A

Fresh Sau3A restriction enzyme was obtained from Biolabs. In order to assay the activity of Sau3A required to give an appropriate degree of partial digestion (Maniatis *et al.*, 1982), 10 μ l of concentrated genomic DNA (approximately 15 μ g) was dissolved in 150 μ l of the buffer recommended by the manufacturer. The DNA was mixed by inverting and the tube was left to sit at room temperature for approximately 1 h. The mixture was then divided into nine aliquots labeled #1 to #9. Tube 1 received 30 μ l, while tubes #2 through #9 received 15 μ l each. Four units (4 U, 1 μ l) of Sau3A were then added to tube #1 and the contents of the tube mixed thoroughly. An aliquot (15 μ l) of this reaction mixture was then transferred to tube #2 and mixed well. An aliquot of tube #2 was then transferred to tube #3, etc. until all but tube #9 had received a dilution of the enzyme. Tube #9 received no enzyme.

Tubes #1 to #9 were incubated at 37 °C for 1 h. At this time, 1 μ l of 0.5 M EDTA and 2 μ l of 10X ficoll load mix (25% ficoll (type 400), 0.25% bromophenol blue, 0.25% xylene cyanol) was added to each tube. The reactions were applied to the wells of a 0.5% agarose gel, and electrophoresed at 100 V for approximately 90 min. The resulting pattern of digestion revealed the amount of enzyme giving the best partial digestion in the range of 10 to 20 kb.

The preparative digest was performed in 3 aliquots, in order to achieve equal representation of sequences cut by the enzyme more or less readily. Forty microlitres of concentrated genomic DNA were diluted to 600 μ l in 1X Sau3A buffer. This was separated into 3 aliquots of 200 μ l. The aliquots were treated with either 1, 2, or 4 U of Sau 3A (4 U/ μ l). The reactions were then incubated at 37 °C for 1 h,

when 50 μ l of 0.5 M EDTA and 70 μ l of 10X ficoll load mix was added. The samples were applied to a low melting point agarose gel (0.6%). When the fragments had separated sufficiently by electrophoresis, the region of the gel containing fragments 10 to 20 kb in length was cut out. The DNA was extracted from the gel slice by phenol extraction as described in the section on basic techniques. Precipitated DNA was resuspended in 150 μ l of TE.

C. Ligations into λ EMBL3 Arms

Before ligations into EMBL3 arms, the genomic DNA fragments resulting from the partial Sau3A digest were tested for their ability to ligate to each other. Samples of the fragments were ligated in a 5 μ l volume. Ligation buffer for BRL T4 DNA ligase (10X) contains 500 mM Tris base, pH 8.0, 70 mM $MgCl_2$, and 10 mM DTT. The test ligations contained 0.5 μ g partially digested DNA, 0.5 μ l 10 mM ATP (freshly diluted from 100 mM), 0.5 μ l 10X ligation buffer, and dH_2O to 5 μ l. One Weiss unit (0.5 μ l) of T4 DNA ligase was used. The reaction mixture was thoroughly mixed and transferred to a siliconized micropipet. Both ends of the pipet were then sealed over a flame. The pipet was incubated overnight in a tube of water at 4 $^{\circ}C$. The degree of ligation was assessed by agarose gel electrophoresis against an unligated control sample.

After determining that self-ligation of the partial digestion fragments was adequate, a preparative reaction was carried out. The EMBL3 arms obtained from the supplier had been digested with 2 enzymes. Digestion with BamHI provided sticky ends homologous to those of the Sau3A fragments. Digestion with EcoRI served to remove the λ stuffer fragment and reduce the probability of stuffer religation. The EMBL3 ligation reactions contained 1 μ l (1 μ g) pre-digested EMBL3 arms, 0.4 μ g

partial digest fragments in 2.5 μ l or less, 0.5 μ l 10 mM ATP, 0.5 μ l 10X ligation buffer, and dH₂O to 4.5 μ l. T4 DNA ligase was added (0.5 μ l, 1 Weiss unit). The mixture was thoroughly mixed and transferred to a siliconized micropipet and sealed as for the test ligation. The pipet was incubated overnight at 4 °C.

D. Packaging of Reconstituted EMBL3

Gigapack packaging extracts were obtained from Vector Cloning Systems Ltd. Extracts were kept frozen at - 70 °C until immediately prior to use. The freeze / thaw extract tube (red) was thawed between the fingers and returned immediately to ice. The sonic extract tube (yellow) was thawed as above and returned to ice. The ligated DNA (4 μ l) was added to the freeze / thaw tube (red). A 15 μ l aliquot of the sonic extract was added to the freeze / thaw extract and mixed gently. The reaction was incubated at room temperature for 2 h. At this time 500 μ l sterile SM (0.1 M NaCl, 8 mM MgCl₂, 50 mM Tris-HCl, pH 7.5, and 0.01% gelatin) and 20 μ l CHCl₃ were added to the tube and the contents mixed well. The reaction was then stored at 4 °C until it was plated onto *E. coli* LE 392 and its P2 lysogen derivative P2 392 in order to determine the titre. The packaging reaction in SM was diluted by factors of 4 X 10², 4 X 10³, and 4 X 10⁴. The phage were adsorbed to host cells for 20 min at 37 °C and plated onto NZ plates.

5. Screening of the Genomic Library

A. Plating the Phage Library

The genomic library prepared above was screened by the procedure of Benton and Davis (1977). The library was initially screened at a high density of 10⁴ plaques per 100 mm petri dish or 5 X 10⁴ plaques per 150 mm petri dish. Aliquots of appropriate

dilutions of the packaging solution were incubated with host cells at 37 °C for 10 min and then plated on NZ plates with the addition of soft NZ agarose. Plates were incubated at 37 °C until the phage plaques were visible but not touching one another. The plates were then incubated at 4 °C for 1 h. Replicas of the plaques were transferred to nitrocellulose circles. In order to allow amplification of the phage plaques, the circles were inverted on fresh NZ plates and incubated at 37 °C overnight. Master plates were stored at 4 °C. For further screens, the amplification step was omitted.

B. Hybridizations and Washing

The DNA on the nitrocellulose filters was denatured by treatment with 0.5 M NaOH, 1.5 M NaCl for 5 min. The pH of the nitrocellulose filters was neutralized by treatment with 1 M Tris-HCl, pH 7.5 for 5 min, followed by 0.5 M Tris-HCl, pH 7.5, 1.5 M NaCl for 5 min. After air drying, the filters were baked at 68 °C for 2 h. Recombinant phage of interest were detected by hybridization to labeled probes and autoradiography. Prior to hybridization, filters were washed with 6X SSC (20X SSC contains 3 M NaCl, 0.4 M Na citrate, pH 7.0), then prehybridized with 6X SSC, 2X Denhardt's (Denhardt's solution contains 0.02% BSA, 0.02% ficoll, and 0.02% polypyrrolidine) at 68 °C for 1 to 4 h. Filters were then hybridized overnight at 68 °C in 6X SSC, 2X Denhardt's, 1 mM EDTA, 0.5% SDS, and denatured, labeled probe. The probe used for screens (pCHIX) was at least 1×10^6 cpm/ml in the hybridization mix, with a specific activity of $>0.5 \times 10^8$ cpm/ μ g. After hybridization, the filters were washed twice at room temperature in 2X SSC followed by three washes at 68 °C in 1X SSC, 0.5% SDS for 30 - 40 min, and finally rinsed in 1X SSC at room temperature. After air drying, the filters were exposed to Kodak XK-1 film overnight at -70 °C with intensifying screens.

When the film was developed, the outlines and orientation markers on the filters were traced onto the film and the plates were oriented using these guides. The appearance of a comet shaped area of exposure revealed the presence of a hybridizing phage clone. For the initial screen, the area of the agar identified by this procedure was lifted out using the large end of a pasteur pipet. For the second and third screens, care was taken to try to isolate as few surrounding phage as possible, so the narrow end of a pasteur pipet was used to isolate the plaques. In both cases the plaques were eluted into 500 μ l of SM in a 1.5 ml microfuge tube. The tube was vigorously vortexed to free the phage from the agar and the mixture was incubated for a minimum of 6 h at 4 °C. After this stock was titred for phage, appropriate dilutions were made and new plates of the isolate were grown. The screening process was then repeated for these new plates.

6. Mapping of the Factor IX EMBL3 Clones

When EMBL3 clones containing regions of the factor IX gene had been isolated, fresh plates were prepared from single, isolated plaques, and stocks of cell lysate were frozen at -70 °C for future use. DNA was isolated from each clone as described in the section on basic techniques.

In order to orient each clone with respect to the map of the factor IX gene, various digestions with restriction enzymes were performed on each of the eight EMBL3 clones isolated during the screen. The restriction enzymes used for these digestions included HindIII, EcoRI, BamHI, SalI, BglII, and KpnI. Southern blots were prepared for several of these gels in order to identify bands containing exon sequence (as described in Part 9). The factor IX cDNA was used as a probe for these blots

(0.8×10^8 cpm/ml). The blots were wetted in 6X SSC. Prehybridization solution contained 6X SSC, 0.5% SDS, 100 µg/ml salmon sperm DNA (denatured), and 2X Denhardt's solution. The blots were prehybridized for 2 h at 68 °C. Hybridization solution contained 6X SSC, 0.5% SDS, 2X Denhardt's, 10 mM EDTA, 100 µg/ml salmon sperm DNA (denatured), and denatured, labeled probe. The blots were hybridized overnight at 68 °C. The first two washes of 10 min each were performed at room temperature in 2X SSC, 0.5% SDS and in 2X SSC, 0.1% SDS, respectively. Three 30 min washes in 0.1X SSC, 0.5% SDS followed. The blots were air dried and exposed to X-ray film overnight with intensifying screens.

7. Subcloning into M13 Vectors for Sequencing

A. Construction of M13 Clones

DNA was sequenced by the chain termination method (Sanger *et al.*, 1977) using the M13 sequencing vectors (Messing *et al.*, 1981; Messing, 1983). DNA to be cloned into M13 vectors was produced by restriction enzyme digestion and recovery of the fragment from low melting point agarose gels (Messing, 1983) or by sonication and end repair (Deininger, 1983). Fragments produced with restriction enzymes were digested under the conditions suggested by the manufacturers. If fragments or mixtures of fragments were to be put directly into a ligation reaction, the digestion mix was heated to 68 °C for 10 min to inactivate the enzymes. The mixture was then extracted with phenol and precipitated from ethanol before ligation. Purified restriction enzyme fragments were isolated from low melting point agarose or polyacrylamide gels.

Random DNA fragments of the factor IX EMBL3 clones were produced by sonication (Deininger, 1983), using a Heat Systems Sonifier at output level 2. DNA

(10 - 20 µg in 0.5 ml of 0.5 M NaCl, 0.1 M Tris-HCl, pH 7.4, 10 mM EDTA) was sonicated in 5 pulses of 5 seconds each. The DNA solution was cooled on ice during the sonication, and was thoroughly mixed between pulses. After the 5 pulses, the DNA was precipitated from solution with ethanol at - 20 °C for 30 min. The precipitated DNA was resuspended into 24 µl of dH₂O. After the addition of 3 µl each of 10X ficoll load mix and 10X TBE, the solution was loaded onto a 5% polyacrylamide non-denaturing gel. The samples were electrophoresed at 50 V for approximately 50 min. After the gel was stained for 10 min in 0.5X TBE with 2 µg/ml of EtBr, the region of the gel containing fragments between 300 and 500 bp was excised from the gel. The DNA contained in the gel fragment was collected by electroelution into 1X TBE at 100 V and precipitated from ethanol at - 70 °C for 1 h. The DNA was resuspended and the solution was extracted 6 times with phenol (1X) and ether (5X). The DNA was again precipitated from ethanol at - 20 °C for 30 min, and was resuspended in 46 µl of 1X S1 buffer (0.25 M NaCl, 30 mM KOAc, 20 mM ZnSO₄, and 10% glycerol).

The ends of the DNA fragments were repaired by incubation with nuclease S1 followed by end repair with the Klenow fragment of DNA Polymerase I. The nuclease S1 reaction consisted of 46 µl DNA solution from above, 4 µl 10X dilute S1 nuclease (5.3 U/µl) for 30 min at room temperature. The reaction was stopped with the addition of Tris base, pH 8.0 to 80 mM and EDTA to 20 mM. The solution was extracted once with a mixture of phenol and chloroform, and precipitated from ethanol at 0 °C for 30 min. The DNA was resuspended in 46 µl of Klenow buffer (20 mM Tris base, pH 8.0, 7 mM MgCl₂). Enzyme (DNA Polymerase 1, Klenow fragment, 7.5 U/µl, 0.1µl used) was added to the reaction and preincubated for 37 °C for 10 min. After 1 µl of each of the dNTPs (10 mM solutions of each) was added,

the reaction was allowed to proceed for another 5 min at 37 ° C. Finally, the reaction mixture was incubated at 68°C for 10 min. The concentration of DNA in the solution was then adjusted to 10 ng/μl with TE.

B. Screening of M13 Clones

On some occasions, a purified fragment was used to ligate into M13 vectors. On other occasions, a mixture of fragments was used. In the latter case, those clones containing exon sequence were identified before sequence analysis. The plaques were screened by plaque hybridization exactly as described for the library screening.

In order to screen for clones containing similar inserts to a clone already isolated, but in the opposite orientation, figure eights between the clones were prepared. An aliquot (1 μl) of DNA from each of two clones to be tested (DNA isolation described below) was added to 14 μl of a solution containing 100 mM Tris-HCl, pH 7.5, 500 mM NaCl, 16% glycerol, 0.2% SDS, and 0.04% bromophenol blue in an 0.5 ml tube. The contents of the tubes were thoroughly vortexed and centrifuged for 1 min in a microfuge. The tubes were incubated for at least 1 h at 68 °C. After the solutions had cooled on ice for 10 min, they were loaded onto 0.7% agarose gels and electrophoresed for 1 h at 100 V. The presence of a clone with a similar insert to the selecting clone, but in the opposite orientation, resulted in the formation of a figure eight structure. This type of complex is retarded in its movement through the gel when compared to the individual clones.

C. M13 DNA Isolation

DNA from clones identified in the previous section was prepared as described by Messing (1983). M13 clones were grown as 2 ml cultures in YT medium in 15 ml

Falcon 2059 tubes using one plaque and 20 μ l of fresh JM103 culture as inoculum. The cultures were incubated at 37 °C for 6 - 16 h. Clones containing larger inserts, > 1 kb, were grown for 6 h. Host cells were removed by centrifugation in a 1.5 ml microfuge tube. Phage particles in 1.3 ml of supernatant were precipitated by the addition of 0.3 ml of 20% PEG, 2.5 M NaCl, and incubation at room temperature for 15 min. M13 phage were collected by centrifugation in an microfuge for 5 min. After removal of all the supernatant, the phage particles were resuspended in 200 μ l of low tris buffer (50 mM NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA). DNA was purified by successive extractions with phenol, phenol:chloroform (1:1, v/v), and chloroform. DNA was precipitated twice from ethanol for 30 min at 0 °C. The final DNA pellet was resuspended in 50 μ l of low tris buffer.

8. DNA Sequencing

DNA in M13 clones was sequenced by the chain termination method (Sanger *et al.*, 1977) as modified for phage M13 templates (Messing *et al.*, 1981). Sequencing reactions were carried out using dideoxy- and deoxy-ribonucleotide solutions of the relative concentrations shown in Table III. Sequencing was performed by hybridizing 4 μ l of template (isolation method in the previous section) with 1 μ l of primer (0.03 OD_{260nm}/ml, 17-mer: 5' -GTAAAACGACGGCCAG- 3'), 1 μ l dH₂O, and 2 μ l 10X Hin buffer (600 mM NaCl, 100 mM Tris-HCl, pH 7.5, 70 mM MgCl₂) at 68 °C for 10 min. After the hybridization mix was allowed to cool to room temperature (20 - 30 min), 1 μ l of 15 μ M dATP and 1.5 μ l of α - ³²P dATP (10 μ Ci/ μ l, 3000 Ci/mmol) were added to the tube. An aliquot (2.5 μ l) of this template / primer mix was mixed with 1.5 μ l of the appropriate deoxy / dideoxy solution (see Table III) in a fresh tube. DNA Polymerase I Klenow fragment (1 μ l of 0.2 U/ μ l) was added. After 15 - 20 min of incubation at room temperature, 1 μ l of

0.5 mM dATP was added. After a further 15 to 20 min at room temperature, 5 μ l of stop / dye mix (98% formamide, 10 mM EDTA pH 8.0, 0.02% xylene cyanole, 0.02% bromophenol blue) was combined with the contents of the tube. The extended products were denatured by heating to 92 °C for 3 min. Aliquots (1 - 2 μ l) of these products were analyzed on 6% and 8% denaturing polyacrylamide gels (50 cm long) at 52 W in 1X TBE. After electrophoresis, the gels were dried at 80 °C under vacuum on a Bio Rad gel drier for 20 - 30 min, and autoradiographed on Kodak XK-1 film overnight at room temperature

Table III.

Composition of dideoxy / deoxy sequencing solutions

Solutions	d/ddG	d/ddA	d/ddT	d/ddC
<u>Nucleotide</u>				
dG	7.9	109.4	158.7	157.9
dT	157.6	109.4	7.9	157.9
dC	157.6	109.4	158.7	10.5
ddG	157.4	-	-	-
ddA	-	116.7	-	-
ddT	-	-	550.3	-
ddC	-	-	-	191.6

The concentrations of the dideoxyribonucleotide (dd) and deoxyribonucleotide (d) triphosphates use in the sequencing solutions for M13 DNA sequencing are shown. Concentrations are given in μ M.

9. Genomic Southern Blot

The genomic Southern blot was transferred to nitrocellulose essentially as described by Southern (1975), and blots were hybridized and washed as described by Kan and Dozy (1978). Genomic DNA (10 µg) was digested with restriction endonucleases (20 - 30 U) in a volume of 40 µl under the conditions suggested by the enzyme manufacturers. DNA fragments were separated by electrophoresis for approximately 16 h at 20 mA in a submerged agarose gel. After electrophoresis, the gel was stained with EtBr so that the DNA in the lanes could be visualized. The gel was exposed to UV light of 260 nm for 90 s to introduce single strand nicks in the DNA, which promote more efficient transfer of the larger fragments. DNA in the gel was then denatured for 45 min in 0.5 N NaOH, 0.6 M NaCl and was then neutralized for 2 periods of 30 min in 1 M Tris-HCl, pH 7.5, 0.6 M NaCl. DNA was transferred to nitrocellulose membranes with 10X SSC for 52 h. After the transfer, the nitrocellulose filter was washed in 3X SSC to remove any agarose, air dried, and then baked at 68 °C for 6 h.

DNA fragments were detected by hybridization to ³²P - labeled probes. The nitrocellulose filter was first wetted with 3X SSC and then prehybridized for 15 h in a solution containing 50% formamide, 3X SSC, 1 mM EDTA, 0.1% SDS, 10 mM Tris-HCl, pH 7.5, 10X Denhardt's solution, 0.05% sodium pyrophosphate, 100 µg/ml denatured herring sperm DNA, 2.5 µg/ml poly(A), and 50 µg/ml tRNA. Hybridizations were carried out in a buffer similar to the above, but lacking the tRNA and containing 6 X 10⁶ cpm/ml of denatured, labeled pCHIX insert. The hybridization reaction was allowed to proceed for 48 h at 37 °C. After hybridization,

the blot was washed for 1 h at room temperature in 2X SSC, 1X Denhardt's, and then washed twice for 90 min at 50 °C in 0.1X SSC, 0.1% SDS. The blot was then rinsed twice at room temperature in 0.1XSSC, 0.1% SDS, followed by four rinses at room temperature in 0.1X SSC. After air drying, blots were exposed to Kodak XK-1 film with intensifying screen for 10 days at - 70 °C.

RESULTS

1. Southern Blot Analysis of the SB IX-4 Factor IX Gene

As a preliminary step in the analysis of the patient's factor IX gene, it was of interest to investigate whether or not gross deletions of the gene had occurred. To address this question, a genomic Southern blot was prepared. A photograph of the autoradiograph from the blot is presented in Figure 3. In this figure, DNA from the patient's father, who is unaffected by hemophilia, was included as a control. The presence of partial digest bands (EcoRI lane, 8.4 kb; BglII lane, 9.5 kb; HindIII lane, 8.5 kb and 14.5 kb) resulted in some differences in intensity between the normal and hemophiliac bands within a given digest. Results from three different restriction enzymes showed that each of the father's factor IX - homologous bands had a similar counterpart in the son's lane. The bands present on the blot were consistent with the fragment sizes predicted by the published factor IX gene map (Yoshitake *et al.*, 1985). These results suggested that no gross deletions or insertions had occurred in the SB IX-4 factor IX gene.

2. Genomic Factor IX Clones Isolated from the SB IX-4 Library

A genomic library of DNA from the patient was then prepared. DNA fragments (10 - 20 kb) from a partial digest of SB IX-4 DNA with the restriction enzyme Sau3A were isolated. The fragments were then introduced into BamHI - digested λ EMBL3 arms. One million independently derived λ clones were screened, using a radioactively

Figure 3
Southern blot analysis of the SB IX-4 factor IX gene

The restriction enzymes BglII, EcoRI and HindIII were used to digest genomic DNA as indicated. Lanes 1,3 and 5: DNA from the unaffected father of the hemophiliac, representing the normal gene. Lanes 2, 4 and 6: SB IX-4 DNA, from the hemophiliac. The normal human factor IX cDNA was used as a hybridization probe. The sizes of selected bands are shown to the right of the blot.

labeled factor IX cDNA as a hybridization probe. Three successive screens of this library yielded eight factor IX containing clones. The map of the restriction enzyme sites of these clones is presented in Figure 4. The EMBL3 clones designated λ IX-12, λ IX-10, and λ IX-6 were selected for further analysis, since together they span the entire gene (Figure 4).

3. M13 Subclones Containing SB IX-4 Factor IX Exons

To facilitate sequence analysis, the DNA inserts from λ clones IX-12, 10, and 6 were fragmented using either restriction enzymes or sonication treatment. These fragments were inserted into M13 vectors. Screens of the subclones with the factor IX cDNA as a hybridization probe allowed identification of those phage inserts containing exons. Positive clones were sequenced using the dideoxy method of Sanger *et al.* (1977). A map of these M13 subclones is presented in Figure 5. Approximately 50% of the coding region sequence was obtained from both coding and non-coding strands. Greater than 80% of the 3' untranslated region was also sequenced. Results from the sequencing of the M13 subclones are presented in Table IV.

4. Nucleotide Sequence of SB IX-4 Activation Peptide and Intron / Exon Junctions

An amino acid polymorphism has been reported in the activation peptide of the factor IX protein (McGraw *et al.*, 1985). Patient SB IX-4 has the ACT codon at amino acid position 3 of the activation peptide (amino acid 148 of the mature protein), as reported by McGraw *et al.* (1985), while others have indicated a GCT codon at this position (Yoshitake *et al.*, 1985) (Table IV). Since the activation peptide is completely removed from factor IX upon activation, this polymorphism is thought to have little effect on the activity of factor IX_a. Both codons have been observed in normal

Figure 4

Genomic factor IX clones isolated from the SB IX-4 library

The lower diagram is a representation of the human factor IX gene (Yoshitake et al., 1985). Exons 1 through 8 are indicated as in Figure 1. EMBL3 clones containing exon sequence were mapped with several restriction enzymes to position the clones relative to the published map of the gene. Each dark line represents a single clone. The sites of cleavage for the restriction enzymes BamHI (B) and EcoRI (E) have been indicated on the line above the gene. The boundaries of the clones extending 3' to the gene were not precisely mapped. Those clones chosen for subcloning are labeled as IX-12, 10 and 6.

Figure 5

M13 subclones containing SB IX-4 factor IX exons

M13 subclones containing exons from the SB IX-4 library are indicated by arrows beneath the exons. Due to the length of the gene (34 kb), only the protein coding regions have been shown. The arrow points in the direction in which the clone was sequenced, from the M13 primer through the clone. Leftward arrows indicate that the coding strand has been sequenced. Rightward arrows indicate that the noncoding strand has been sequenced. Each clone is labeled with its name. Only the translated region has been presented, although clones containing 3' untranslated region were also obtained. Restriction enzyme sites used in cloning are indicated below the exons. Scale: |-----| 100 bp.

Table IV

Nucleotide sequences of the SB IX-4 activation peptide and intron / exon junctions

----- ACTIVATION PEPTIDE -----			
NH ₂ - - AAG CTC ACC CGT / GCT GAG GCT GTT TTT CCT - - COOH			
lys	leu	thr	arg
			145
			146
			ala
			glu
			ala
			val
			phe
			pro
- - AAG CTC ACC CGT / GCT GAG ACT GTT TTT CCT - -			
			thr
----- INTRON / EXON JUNCTIONS -----			
intron	exon / 5'	intron	3'/exon
A	CAG/GTTTGT.....	TTTCAG/T	
B	ACA/GTGAGT.....	TTATAG/A	
C	TTG/GTAAGC.....	TCAAAG/A	
D	TAG/GTAAGT.....	TTTTAG/A	
E	CAG/GTCATA.....	TTCTAG/T	
F	CAG/GTACTT.....	TCACAG/G	
G	CAG/GTAAAT.....	TAATAG/G	
consensus	CAG/GT A AGT.....	TT N TAG/G	
	A G	CC C	
after Yoshitake <i>et al.</i> , 1985			

In the first section, the polymorphic activation peptide residue is shown. The N-terminal activation cleavage site, between amino acids 145 and 146, is shown by a slash. The polymorphic amino acid is shown in bold type: the upper line represents the most commonly observed sequence, and the lower line the sequence found by McGraw *et al.*, 1985 and in SB IX-4. The lower table lists the sequences found at the factor IX intron / exon junctions and lists the consensus sequences identified by Mount, 1982, and Sharp, 1981. In each case the slash represents the site of splicing. All intron / exon junctions sequences in SB IX-4 were identical to those identified by Yoshitake *et al.*, 1985.

individuals. Knowledge of this polymorphism has been exploited in carrier detection studies using an oligonucleotide probe (Winship and Brownlee, 1986) and an antibody (Wallmark *et al.*, 1985) specific to this region.

In addition to the coding regions, flanking portions of introns and intron / exon junctions were compared to the normal sequence. Table IV notes the sequences observed in these areas. All intron / exon boundaries displayed the invariant GT / AG dinucleotides associated with introns (Breathnach and Chambon, 1981). The sequence of the introns extending from such junctions was determined for a minimum of 25 base pairs. In each case the sequence observed was identical to that reported by Yoshitake *et al.* (1985). The putative sites for transcript cleavage and poly(A) addition were sequenced and were found to be identical to the sequence presented by Yoshitake *et al.* (1985). The putative promoter region was found to be identical to the normal sequence for 225 base pairs upstream of exon 1.

5. Nucleotide Sequence Alteration in the SB IX-4 Factor IX Gene

One nucleotide change relative to the normal protein coding sequence was observed at nucleotide 31,311 of the factor IX gene (numbering system of Yoshitake *et al.*, 1985). This corresponds to amino acid 397 of the mature protein, which is found in the serine protease subunit of factor IX_a. The normal isoleucine codon, ATA, has been changed to a threonine codon, ACA. This sequence has been observed in two independently obtained non-coding strand clones (8C-5 and 8C-17), and from one coding strand clone derived by reversing the insert of 8C-5. Figure 6 shows photographs of the sequencing gel autoradiographs. The DNA sequence of this region and the nucleotide change relative to the normal sequence are shown.

Figure 6

Nucleotide sequence alteration in SB IX-4 factor IX gene

The sequence of the alteration at nucleotide 31,311, corresponding to amino acid 397, is given for both the coding and noncoding strands of DNA. The sequence of interest is given to the left of the autoradiogram and the horizontal arrow points directly toward the mutation. The amino acid sequences given in the boxes show the direction of sequencing. The direction of the arrow indicates the way the sequence is to be read from the bottom to the top of the autoradiogram.

CODING STRAND

normal		SB IX-4	
A		A	
	T		C
A		A	
A		A

NONCODING STRAND

normal		SB IX-4	
	T		T
	T		T
A		G	
	T		T

DISCUSSION

1. Summary of Data Concerning Patient SB IX-4

A significant array of data have now been compiled about the molecular defect causing decreased factor IX clotting activity in the patient SB IX-4. Antigen and activity assay results indicate a significant reduction in factor IX_a activity with respect to factor IX antigen levels. A Southern blot has suggested that no major deletions of the factor IX gene are present. All eight exons were isolated from the genomic library prepared from the patient's DNA. Within the coding regions, a single nucleotide change relative to the normal sequence (Yoshitake *et al.*, 1985) was observed. This change results in the substitution of a threonine codon, ACA, for the normal isoleucine codon, ATA. It should result in an amino acid alteration at residue 397 of the protein. These results are consistent with each other and suggest that the observed nucleotide change is the cause of the decrease in specific activity in the factor IX of SB IX-4.

2. Naming of the Mutation Found in SB IX-4

No other mutation in the serine protease region of factor IX has been identified at the molecular level. In keeping with the tradition of naming a newly identified mutation of factor IX by a subscript naming the place of birth of the patient, we have given the name factor IX_{Vancouver} to the abnormal protein found in patient SB IX-4.

3. Implications for Carrier Detection and Prenatal Diagnosis

Several restriction fragment length polymorphisms (RFLP's) have been identified in the factor IX gene (Camerino *et al.*, 1985; Camerino *et al.*, 1984; Winship *et al.*, 1984; Hay *et al.*, 1986; Giannelli *et al.*, 1984). These polymorphisms have been

used extensively in the diagnosis of hemophilia B carrier status and in prenatal diagnosis. Due to linkage disequilibrium between several of these RFLP sites, however, diagnosis is not possible in all cases.

Several studies have shown that knowledge of a point mutation can be used in the diagnosis of human diseases. Such diagnoses were first obtained for sickle cell anemia (Conner *et al.*, 1983). Recently, the method has been applied to hemophilia B (Rees *et al.*, 1985; Winship and Brownlee, 1986). In these studies, pairs of 19 base long oligonucleotide probes were prepared. One probe contained a region of the normal sequence, and the other contained the same region but the abnormal or polymorphic sequence. By careful manipulation of hybridization and washing conditions, it was possible to distinguish between the normal and mutant alleles on Southern blots.

In the case of factor IX_{Vancouver}, oligonucleotide probes could be synthesized which match the normal and mutant sequences between amino acids 394 to 400 (nucleotides 31,302 to 31,320). These probes could be used in carrier detection and prenatal diagnosis. Since the proportion of hemophilia B variants involving a mutation at isoleucine 397 may be small, this analysis might be of limited general usefulness. However, if a series of such specific probes were to become available, it might be possible to obtain molecular diagnoses for many variants of hemophilia B.

4. Possible Effects of the Ile₃₉₇ to Thr₃₉₇ Substitution on Factor IX_a Activity:

A. Structure of Normal Factor IXa

The three dimensional structures of factor IX_a and its substrate, factor X, have not been determined by X-ray crystallography since, like many of the blood clotting factors, these proteins have not been crystallized. Therefore, the role of isoleucine 397 in the activity of normal factor IX_a is not known. Factor IX_a activates the next clotting factor in the cascade, factor X, by cleaving the peptide bond between arginine₁₉₄ and alanine₁₉₅. The precise points of interaction of the proteins in the factor IX_a - factor X complex, however, are not known.

B. Structural Homology Between the Trypsin-Like Serine Proteases

The term 'serine protease' is applied to a group of structurally related endoproteases which have a serine residue at the active site. These enzymes have been isolated from a variety of organisms. On the basis of amino acid sequence homologies, serine proteases are divided into two major families, the subtilisins and the trypsin-like enzymes. Factor IX is a member of the trypsin-like family. Trypsin is one of several well known serine proteases which are involved in the process of digestion. Two other digestive proteases, α chymotrypsin and elastase, have also been studied in detail. The three dimensional structures of all three of these enzymes have been determined by X-ray crystallography.

An alignment of the amino acid sequences of these three serine proteases with the sequence of factor IX_a (Table V), shows the striking similarity between these enzymes. Since it is traditionally applied to all the enzymes in order to simplify comparisons (Hartley and Shotton, 1971), the chymotrypsin numbering system has

Table V

Amino acid sequence alignment of normal human factor IX_a
and the pancreatic serine proteases

The amino acid sequences of the catalytic regions of normal human factor IX_a, bovine trypsin (TRYP), porcine elastase (ELAS) and bovine chymotrypsin (CHYM) are shown. The sequences were taken from Yoshitake *et al.* (1985), Marquart *et al.* (1983), Sawyer *et al.* (1978) and Cohen *et al.* (1981), respectively. The numbering system shown here is based on the system developed for chymotrypsinogen A (Hartley and Kauffman, 1966), with insertions in the sequences of related enzymes denoted by letters (i.e. 61a, 61b). Deletions are indicated by dashes. Identical residues in corresponding positions of all four proteins are boxed. The standard single letter code for amino acids is used.

TRYP	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	36	36	36
ELAS	I	V	G	G	Y	T	C	G	A	N	T	V	P	Y	Q	V	S	L	N	-	R	S	G	S
CHYM	I	V	N	G	E	E	A	V	P	G	S	W	P	S	Q	I	S	L	Q	Y	K	-	-	
FIXa	V	V	G	G	E	D	A	K	P	G	Q	F	P	W	Q	V	V	L	N	G	K	-	-	
TRYP	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	
ELAS	S	G	Y	H	F	C	G	G	S	L	I	N	S	Q	W	V	V	S	A	A	H	C	Y	
CHYM	T	G	F	H	F	C	G	G	S	L	I	N	E	N	W	V	V	T	A	A	H	C	V	
FIXa	V	D	-	A	F	C	G	G	S	I	V	N	E	K	W	I	V	T	A	A	H	C	F	
TRYP	60	a	b		62	63	a	b	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	
ELAS	K	S	-	G	I	Q	V	R	V	-	-	G	E	D	N	L	N	Q	V	N	G	T	E	
CHYM	D	R	-	E	K	T	F	R	V	V	V	G	E	H	N	I	N	S	N	N	E	K	I	
FIXa	V	T	G	T	S	D	V	-	V	V	A	G	E	F	D	Q	E	S	S	E	H	T	E	
TRYP	81	82	83	84	85	86	87	88	89	90	91	92	93	94	a	b	c	96	97	98	a	b	c	
ELAS	Q	F	I	S	A	S	K	S	I	V	H	P	S	Y	N	-	-	S	N	T	L	-	-	
CHYM	Q	Y	V	G	V	Q	K	S	V	V	H	P	Y	W	N	-	-	T	D	D	V	A	-	
FIXa	Q	K	R	N	I	I	K	I	V	K	N	S	K	Y	N	A	-	I	L	T	I	-	-	
TRYP	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121		
ELAS	N	N	D	I	M	L	I	K	L	K	S	A	A	S	L	N	S	R	V	A	S	I		
CHYM	G	Y	D	I	A	L	L	R	L	A	Q	S	V	A	F	N	S	Y	V	Q	L	G		
FIXa	N	H	D	I	A	L	L	E	L	D	E	P	L	V	L	N	S	Y	V	T	P	I		
TRYP	122	123	124	125	a	b	c	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	
ELAS	S	L	P	T	-	-	-	S	C	A	S	-	A	G	T	Q	C	L	I	S	G	W	G	
CHYM	V	L	P	R	A	-	-	G	T	I	A	-	A	N	T	P	C	Y	I	T	G	W	G	
FIXa	C	I	M	D	K	E	Y	T	N	I	F	L	K	F	G	S	G	Y	V	S	G	W	G	
TRYP	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	
ELAS	N	T	K	S	S	G	T	S	Y	P	D	T	L	K	C	L	K	A	P	I	L	S	N	
CHYM	L	T	R	-	T	N	G	Q	L	P	Q	V	L	Q	Q	A	Y	S	P	T	V	D	Y	
FIXa	R	V	F	H	-	K	G	R	S	A	L	V	L	Q	Y	L	R	V	P	L	V	D	R	
TRYP	166	167	168	169	a	b	c	171	172	173	174	175	176	177	178	179	180	181	182	183	a	b	185	
ELAS	S	S	C	K	170	170	170	Y	-	P	G	Q	I	T	S	N	M	F	C	A	G	-	L	
CHYM	A	I	C	S	S	S	S	Y	W	G	G	T	V	K	N	S	M	V	C	A	G	-	G	
FIXa	A	T	C	L	R	-	-	S	T	K	F	T	I	Y	N	N	M	F	C	A	G	-	F	
TRYP	a	b	c	187	a	b	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205
ELAS	E	-	-	G	G	G	K	D	S	C	Q	G	D	S	G	G	P	V	V	C	S	-	-	G
CHYM	N	-	-	G	V	V	R	S	S	C	Q	G	D	S	G	G	P	L	H	V	K	V	N	G
FIXa	H	E	G	G	R	-	-	D	S	C	Q	G	D	S	G	G	P	H	V	T	E	V	E	G
TRYP	206	207	208	209	210	211	212	213	214	215	216	a	b	218	219	220	a	b	222	223	224	225	226	
ELAS	-	G	K	L	Q	G	I	V	S	W	G	S	-	-	G	C	A	Q	K	N	K	P	G	
CHYM	A	Y	A	V	H	G	V	T	S	F	V	S	R	-	T	C	N	V	T	R	K	P	T	
FIXa	T	S	F	L	T	G	I	I	S	W	G	E	-	-	E	C	A	M	K	G	K	Y	G	
TRYP	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245					
ELAS	V	Y	T	K	V	C	N	Y	V	S	W	I	K	Q	T	I	A	S	N					
CHYM	V	F	T	R	V	A	A	Y	I	S	W	I	N	N	V	I	A	S	N					
FIXa	I	Y	T	K	V	S	R	Y	V	N	W	I	K	E	K	T	K	L	T					

been used. Insertions and deletions relative to the chymotrypsin sequence have been accommodated by giving the same number but different letters to each residue in the inserted regions. Those positions at which there is identity between the residues of all four of the proteins have been enclosed in boxes.

When the amino acid sequences of the pancreatic serine proteases are compared to one another, there is approximately 41% identity of amino acids between them (James *et al.*, 1978). When the known three dimensional structures of the pancreatic enzymes are compared by superposition of the backbone chains, however, an even greater similarity is evident (~ 85% of the α - carbon atoms are topologically equivalent; James *et al.*, 1978). The levels of primary sequence identity among these enzymes (~ 41%), and between them and human factor IX_a (~ 36%), are similar. Therefore it is reasonable to expect a corresponding level of tertiary structure similarity between factor IX_a and the pancreatic serine proteases.

In addition to the structural similarity of the proteases, the three dimensional structures of portions of several of the known inhibitors of serine proteases are similar. Bovine pancreatic trypsin inhibitor (BPTI) and soy bean trypsin inhibitor (SBTI) are two well known examples. Those regions of the inhibitors which interact directly with the active site show as much structural similarity between them as do the structures of the proteases themselves (Chen and Bode, 1983). While the binding in the complexes formed between these inhibitors and the proteases is thought to be stronger than that between the proteases and their substrates, they have been shown to react with the enzymes by the same pathway. The overall structure of the two kinds of complexes is thought to be similar (Chen and Bode, 1983; and references therein).

C. Model of the Complex Between Factor IX_a and Its Substrate Peptide

The similarity with the pancreatic serine proteases and others, such as kallikrein, can be exploited to prepare models of the approximate structure of the blood clotting serine proteases. Building on the known coordinates of the polypeptide chain atoms of the pancreatic enzymes, several models of the three dimensional structure of human factor IX_a have been prepared (Furie *et al.*, 1985; Greer, 1984; and G. Louie and G. Brayer, personal communication). Models have also been developed for prothrombin, factor X and factor XII (Furie *et al.*, 1982; Cool *et al.*, 1985). These models have allowed identification of regions of conserved structure and regions of variability between the enzymes (Furie *et al.*, 1982). While the cores of the molecules appear to be highly conserved, the surface features are variable between the enzymes. The blood clotting serine proteases are highly specific for their protein substrates. Since the active sites of these enzymes are nearly identical, it is the protein-protein interactions outside the direct substrate binding regions which are thought to allow specific recognition of the protein substrate (Craik *et al.*, 1985).

In addition to the models of factor IX_a, a model of the structure of an eight residue peptide containing the activation cleavage sequence of factor X has been prepared. This model is based on the conformation of that portion of the bovine pancreatic trypsin inhibitor which is bound in the active site of trypsin (Louie and Brayer, personal communication). The bond cleaved by factor IX_a is part of the sequence N⁻₁₉₁asn-leu-thr-arg==ile-val-gly-gly₁₉₈-C in the 55 kDa factor X protein (Fung *et al.*, 1984). The model of this substrate peptide can be inserted into the model of factor IX_a. The resulting complex will be referred to as the factor IX_a - substrate complex. Using this model, interactions between these proteins can be studied.

The model of the factor IX_a - substrate complex prepared by Louie and Brayer (personal communication) has been analyzed with respect to the position of the factor IX_{Vancouver} mutation at residue 397. In the chymotrypsin numbering system, amino acid 397 is referred to as amino acid 227. As shown in the three dimensional model of factor IX_a (Figure 7A), isoleucine 227 (IX 397) occupies a position adjacent to the active site of the enzyme.

Five structural features of serine proteases are thought to be crucial to their unique reactivity (Kraut, 1977). The active site serine is in a position to form a covalent bond with the carbonyl carbon atom of the susceptible bond. There is a site (termed the 'oxyanion hole') for binding of the carbonyl oxygen atom when the carbonyl is in the tetrahedral transition state. An S1 binding pocket accommodates the side chain of the amino acid on the acyl side of the bond to be cleaved and thus provides substrate specificity. On the amino side of the susceptible bond, the polypeptide chain of the substrate is weakly bound to the enzyme by hydrogen bonds. The weakness of this binding facilitates the removal of the first product of the reaction from the active site. Finally, there is extensive binding between that portion of the substrate polypeptide chain acyl to the susceptible bond and the surface of the serine protease.

Residue 397 of factor IX_a is located near the surface involved in this extensive hydrogen bonding with the acyl portion of the bound substrate. In the region of factor IX_a around residue 397, therefore, there are believed to be significant interactions between factor IX_a and its substrate, factor X (Chen and Bode, 1983). By analogy to the bonding of the other serine proteases with their various inhibitors, these interactions would be expected to involve hydrogen bonding and / or hydrophobic interactions (Kraut 1977).

Figure 7

Hydrogen bonding in the factor IX_a - substrate complex models

Photographs of the computer screen show the three dimensional structure of the factor IX_a model's active site (Louie and Brayer, personal communication). The catalytic triad residues are shown in red and are labelled with their names and amino acid number. They are serine 195, histidine 57 and aspartate 102. The side chains of isoleucine 227 (IX 397) of the normal molecule and threonine 227 (IX 397) of factor IX_{Vancouver} are drawn in blue. The relevant hydrogen bonds are indicated as dashed yellow lines. The numbers on these lines refers to the interatomic distance, in Angstroms. The atoms involved in the hydrogen bonds have been labeled in white. The factor X substrate peptide has been indicated in green.

- A) The position of ile 227 (IX 397) relative to the active site residues in the normal factor IX_a model.
- B) Two of the potential hydrogen bonds in the model of the complex formed between normal factor IX_a and the factor X substrate peptide are shown.
- C) The potential hydrogen bond unique to factor IX_{Vancouver} is indicated. The formation of this bond may interrupt the normal hydrogen bonding shown in part B.

A.

B.

C.

D. Hydrogen Bonding in the Normal Factor IX_a -Substrate Complex

Examination of the model of the normal factor IX_a - factor X peptide complex reveals several potential sites for hydrogen bonding. The presence of donor and acceptor ligands, an appropriate geometry for interaction, and the close proximity (~ 2.5 to 3.5Å) of these ligands to one another are the factors which determine the possibility of hydrogen bonding. One of the possible interchain hydrogen bonds in the normal factor IX_a - factor X peptide complex appears between the amide nitrogen of glycine 216 (IX 386) of factor IX_a and the carbonyl oxygen of leucine 192 (also called leucine 13) of the factor X peptide (Figure 7B).

E. Hydrogen Bonding in the Factor IX_a Vancouver - Substrate Complex

From the model of the factor IX_{Vancouver} - substrate complex, it appears that an additional intramolecular hydrogen bond could form between two factor IX_a atoms. An extra bond is possible since the non-hydrogen bonding side chain of isoleucine 397 is replaced by a hydroxyl group in the threonine side chain. The additional hydrogen bond could form between the hydroxyl group on the side chain of threonine 227 (IX 397) and the carbonyl oxygen of the peptide bond of tryptophan 215 (IX 385) (Figure 7C). If this abnormal intrachain bond was to form, the normal hydrogen bonding between leucine 192 of factor X and glycine 216 (IX 386) of factor IX_a (described above) could be disrupted. It is possible that such a change in the hydrogen bonding between the enzyme and its substrate could lead to a less productive orientation of the scissile bond of factor X in the active site of factor IX_a. This could result in a significantly reduced rate of factor IX_a catalytic activity. Thus, the three dimensional model of factor IX_a, although approximate, has provided a possible rationale for the observed decrease in specific coagulant activity associated with factor IX_{Vancouver}.

5. Directions for Further Research

Further studies on the structure of factor IX_{Vancouver} are possible. The patient under study has sufficient antigen in his plasma to allow the purification of factor IX_{Vancouver}. Protein analysis could be used to compare characteristics of factor IX_{Vancouver} with normal factor IX: calcium binding could be quantitated, activation cleavages could be performed and compared to those of normal factor IX, and the molecular weight of the mutant protein could be estimated (Bertina and van der Linden, 1982b).

A more complete understanding of the effects of the isoleucine to threonine mutation at amino acid 397 may be more difficult to obtain. Since the mutation in factor IX_{Vancouver} is thought to affect binding of factor X to factor IX_a, a direct test of the mutation would involve some form of kinetic measurements. Coagulation assays measure the rate of formation of a fibrin clot. Since a great many enzymes and cofactors are involved in the formation of a clot, coagulation assays are heterogeneous, difficult to standardize, and unsuitable for kinetic studies of the individual enzymes (Stormorken, 1979).

Kinetic experiments on isolated factor IX_{Vancouver} may be hampered by the fact that highly reactive *in vitro* substrates for factor IX_a are not yet available. A number of synthetic substrates for the blood clotting serine proteases have been prepared. Substrates tested for activity with bovine factor IX_a include ethyl ester, thioester, 4 - nitroanilide, and fluorogenic peptides and amino acids (Lindquist *et al.*, 1978; Byrne and Castellino, 1978; Byrne *et al.*, 1980; McRae *et al.*, 1981; Link and Castellino, 1983; Castillo *et al.*, 1983; and Cho *et al.*, 1984). Factor IX_a has been

identified as the least reactive and most specific of the blood clotting serine proteases with regard to these synthetic substrates. In addition to the problem of low reactivity of the substrates, their limited size may hinder demonstration of altered substrate binding by factor IX_{Vancouver}. It is possible that less direct, two stage assays, in which the activation of factor X by factor IX_a is followed with chromogenic substrates of factor X_a, could be used to study the effect of the factor IX_{Vancouver} mutation (Griffith *et al.*, 1982; Suomela *et al.*, 1977). Several synthetic and natural inhibitors of factor IX_a have been reported (Rosenberg *et al.*, 1975; Kurachi *et al.*, 1976; Lollar and Fass, 1984; and Turner *et al.*, 1986) which can be used in the identification or quantitation of the factor IX_a active site.

A detailed analysis of the interaction of factor IX_{Vancouver} with its substrate may thus not be feasible until more reactive synthetic substrates are prepared or until factor IX_a is crystallized for X-ray diffraction studies. Indirect lines of investigation, however, are open. This mutation will be of interest to enzymologists due to the fact that few mutations in serine proteases have been reported. Site directed mutagenesis of the homologous amino acid in trypsin, chymotrypsin or elastase may allow testing of the possibility of altered substrate binding. Since it is easier to solve the structures of molecules differing by only one residue from a known structure, the exact structures of the mutant molecules could be determined. Kinetic studies of normal and altered molecules could be performed if suitable substrates are available.

6. Summary

A nucleotide alteration has been identified in the factor IX gene isolated from patient SB IX-4. The alteration changes an amino acid in the serine protease subunit of factor IX_a. Since this is the only observed departure from the normal sequence of the protein coding region of the gene, it is likely to be the molecular defect responsible for the decrease in specific coagulant activity of factor IX in this hemophiliac. A three dimensional model of human factor IX_a, which is based on its homology to the pancreatic serine proteases, shows that the altered amino acid is very close to the active site. The change from an isoleucine side chain, which cannot participate in hydrogen bonding, to a threonine side chain, which carries a hydrogen bonding ligand, could disrupt the normal hydrogen bonding between factor IX and its substrate, factor X. This disruption could result in the reduced specific activity observed in the SB IX-4 factor IX_a molecule. In accordance with standard practice, the mutation has been given the name factor IX_{Vancouver}.

REFERENCES

- Anson, D.S., Choo, K.H., Rees, D.J.G., Giannelli, F., Gould, K., Huddleston, J.A. and Brownlee, G.G. (1984). "The gene structure of human anti-haemophilic factor IX." *EMBO J.* **3**: 1053-1060.
- Bajaj, S.P. (1982). "Cooperative Ca^{2+} Binding to Human Factor IX." *J. Biol. Chem.* **257**: 4127-4132.
- Bell, G.I., Karam, J.H., and Rutter, W.J. (1981). "Polymorphic DNA region adjacent to the 5' end of the human insulin gene." *Proc. Natl. Acad. Sci. USA* **78**: 5759-5763.
- Bentley, A.K., Rees, D.J.G., Rizza, C., and Brownlee, G.G. (1986). "Defective Propeptide Processing of Blood Coagulation Factor IX Caused by Mutation of Arginine to Glutamine at Position - 4." *Cell* **45**: 343-348.
- Benton, W.D. and Davis, R.W. (1977). "Screening λ gt Recombinant Clones by Hybridization in situ." *Science* **196**: 180-182.
- Berget, S.M. (1984). "Are U4 Small Nuclear Ribonucleoproteins at the 5' terminus of Adenovirus 2 late mRNA?" *Proc. Natl. Acad. Sci. USA* **74**: 1371-1375.
- Bertina, R.M. and van der Linden, I.K. (1982a). "Factor IX Zutphen: A genetic variant of blood coagulation factor IX with an abnormally high molecular weight." *J. Lab.Clin. Med.* **100**: 695-704.
- Bertina, R.M. and van der Linden, I.K. (1982b). "Detection and Classification of Molecular Variants of Factor IX." in Bloom, A.L. (ed): "The Hemophilias." Churchill Livingstone, New York. pp. 137-155.
- Birnboim, H.C. and Doly, J. (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." *Nucleic Acids Res.* **7**: 1513.
- Birnstiel, M.L., Busslinger, M., and Strub, K. (1985). "Transcription Termination and 3' Processing : The End is in Site." *Cell* **41**: 349-359.
- Blomquist, M.C., Hunt, L.T., and Barker, W.C. (1984). "Vaccinia virus 19-kilodalton protein: Relationship to several mammalian proteins, including two growth factors." *Proc. Natl. Acad. Sci. USA* **81**: 7363-7367.
- Boyd, Y., Buckle, V.J., Munro, E.A., Choo, K.H., Migeon, B.R., and Craig, I.W. (1984). "Assignment of the haemophilia B (Factor IX) locus to the q26-qter region of the X chromosome" *Ann. Hum. Genet.* **48**: 145-152.
- Bray, G.L., and Thompson, A.R. (1986). "Partial Factor IX Protein in a Pedigree with Hemophilia B Due to a Partial Gene Deletion." *J. Clin. Invest.* **77**: 1194-1200.

- Bray, G.L., Weinmann, A.F., and Thompson, A.R. (1986). "Calcium-specific immunoassays for factor IX: Reduced levels of antigen in patients with vitamin K disorders." *J. Lab. Clin. Med.* **107**: 269-278.
- Breathnach, R., and Chambon, P. (1981). "Organization and Expression of Eucaryotic Split Genes Coding for Proteins." *Ann. Rev. Biochem.* **50**: 349-383.
- Briet, E., Bertina, R.M., van Tilburg, N.H., and Veltkamp, J.J. (1982). "Hemophilia B Leyden: A Sex-Linked Hereditary Disorder That Improves after Puberty." *N. Engl. J. Med.* **306**: 788-790.
- Brown, P.E., Hougie, C., and Roberts, H.R. (1970). "The genetic heterogeneity of hemophilia B." *N. Engl. J. Med.* **283**: 61.
- Byrne, R. and Castellino, F.J. (1978). "The Influence of Metal Ions in the Activation of Bovine Factor IX by the Coagulant Protein of Russell's Viper Venom." *Arch. Biochem. Biophys.* **190**: 687-692.
- Byrne, R., Link, R.P., and Castellino, F.J. (1980). "A Kinetic Evaluation of Activated Bovine Blood Coagulation Factor IX toward Synthetic Substrates." *J. Biol. Chem.* **255**: 5336-5341.
- Camerino, G., Oberle, I., Drayna, D., and Mandel, J.L. (1985). "A New MspI restriction fragment length polymorphism in the hemophilia B locus." *Hum. Genet.* **71**: 79-81.
- Castillo, M.J., Kurachi, K., Nishino, N., Chkubo, I. and Powers, J.C. (1983). "Reactivity of Bovine Blood Clotting Factor IX_{aβ}, Factor X_{aβ}, and Factor XI_a toward Fluorogenic Peptides Containing the Activation Site Sequences of Bovine Factor IX and Factor X." *Bioc.* **22**: 1021-1029.
- Chen, Z., and Bode, W. (1983). "Refined 2.5Å X-ray Crystal Structure of the Complex Formed by Porcine Kallikrein A and the Bovine Pancreatic Trypsin Inhibitor." *J. Mol. Biol.* **104**: 283-311.
- Cho, K., Tanaka, T., Cook, R.R., Kisiel, W., Fujikawa, K., Kurachi, K. and Powers, J.C. (1984). "Active-Site Mapping of Bovine and Human Blood Clotting Serine Proteases Using Synthetic Peptide 4 - Nitroanilide and Thio Ester Substrates." *Bioc.* **23**: 644-650.
- Choo, K.H., Gould, K.G., Rees, D.J.G., and Brownlee, G.G. (1982). "Molecular cloning of the gene for human anti-haemophilic factor IX." *Nature* **299**: 178-180.
- Cohen, G.H., Silverton, E.W., and Davies, D.R. (1981). "Refined Crystal Structure of γ - Chymotrypsin at 1.9Å Resolution." *J. Mol. Biol.* **148**: 449-479.

- Conner, B.J., Reyes, A.A., Morin, C., Ikatura, K., Teplitz, R.L., and Wallace, R.B. (1983). "Detection of sickle cell β^S - globin allele by hybridisation with synthetic oligonucleotides." *Proc. Natl. Acad. Sci. USA* **80**: 278-282.
- Cool, D.E., Edgell, C-J., Louie, G.V., Zoller, M.J., Brayer, G.D., and MacGillivray, R.T.A. (1985). "Characterization of Human Blood Coagulation Factor XII cDNA." *J. Biol. Chem.* **260**: 13666-13676.
- Craik, C.S., Largman, C., Fletcher, T., Rocznik, S., Barr, P.J., Fletterick, R., and Rutter, W.J. (1985). "Redesigning Trypsin: Alteration of Substrate Specificity." *Science* **228**: 291-297.
- Dagert, M. and Ehrlich, S.D. (1979). "Prolonged Incubation in Calcium Chloride Improves the Competence of *Escherichia coli* Cells." *Gene* **6**: 23-28.
- Davis, L.M., McGraw, R.A., Graham, J.B., Roberts, H.R., and Stafford, D.W. (1984). "Identification of the genetic defect in factor IX Alabama: DNA sequence analysis reveals a Gly substitution for Asp⁴⁷." *Blood* **64**: 262a (Suppl 1) (abstr.).
- Deininger, P.L. (1983). "Random Subcloning of Sonicated DNA: Application to Shotgun DNA Sequence Analysis." *Anal. Biochem.* **129**: 216-223.
- Doolittle, R.F., Feng, D.F., and Johnson, M.S. (1984). "Computer-based characterization of epidermal growth factor precursor." *Nature* **307**: 558-560.
- Diuguid, D.L., Rabin, M.J., Furie, B.C., Liebman, H.A., and Furie, B. (1986). "Molecular basis of hemophilia B: A defective enzyme due to an unprocessed propeptide is caused by a point mutation in the factor IX precursor." *Proc. Natl. Acad. Sci. USA* **83**: 5803-5807.
- Enfield, D.L., and Thompson, A.R. (1984). "Cleavage and Activation of Human Factor IX by Serine Proteases." *Blood* **64**: 821-831.
- Fantl, P., Sawers, R.J., and Marr, A.G. (1956). "Investigation of a haemorrhagic disease due to β prothromboplastin deficiency complicated by a specific inhibitor of thromboplastin formation." *Aust. Ann. Med.* **5**: 163.
- Feinberg, A.P. and Vogelstein, B. (1983). "A Technique for Radiolabeling DNA Restriction Endonuclease Fragments to High Specific Activity." *Anal. Biochem.* **132**: 6-13.
- Frischauf, A-M., Lehrach, H., Poutska, A., and Murray, N. (1983). "Lambda Replacement Vectors Carrying Polylinker Sequences." *J. Mol. Biol.* **170**: 827-842.
- Fujikawa, K., Coan, M.H., Legaz, M.E., and Davie, E.W. (1974). "The Mechanism of Activation of Bovine Factor X (Stuart Factor) by Intrinsic and Extrinsic Pathways." *Biochemistry* **13**: 5290-5299.

- Fung, M.R., Campbell, R.M., and MacGillivray, R.T.A. (1984). "Blood coagulation factor X mRNA encodes a single polypeptide chain containing a prepro leader sequence." *Nucleic Acids Res.* **12**: 4481-4492.
- Fung, M.R., Hay, C.W., and MacGillivray, R.T.A. (1985). "Characterization of an almost full-length cDNA coding for human blood coagulation factor X." *Proc. Natl. Acad. Sci. USA* **82**: 3591-3595.
- Furie, B., Bing, D.H., Feldman, R.J., Robison, D.J., Burnier, J.P., and Furie, B.C. (1982). "Computer-Generated Models of Blood Coagulation Factor X_a, Factor IX_a, and Thrombin Based on Structural Homology with Other Serine Proteases." *J. Biol. Chem.* **257**: 3875- 3882.
- Giannelli, F., Choo, K.H., Rees, D.J.G., Boyd, Y., Rizza, C.R., and Brownlee, G.G. (1983). "Gene deletions in patients with haemophilia B and anti-factor IX antibodies." *Nature* **303**: 181-182.
- Giannelli, F., Choo, K.H., Winship, P.R., Rizza, C.R., Anson, D.S., Rees, D.J.G., Ferrari, N., and Brownlee, G.G. (1984). "Characterisation and use of an intragenic polymorphic marker for detection of carriers of haemophilia B (factor IX deficiency)." *Lancet* **i**: 239-241.
- Gray, C.R., Teasdale, J.M. and Thomas, J.W. (1968). "Hemophilia B_M." *Canad. Med. Ass. J.* **98**: 552-554.
- Greer, J. (1985). "Protein Structure and Function by Comparative Model Building." *Ann. N.Y. Acad.Sci.* **439**: 44-63.
- Griffith, M.J., Reimer, H.M., Lundblad, R.L. and Roberts, H.R. (1982). "Measurement of Human Factor IX_a Activity in an Isolated Factor X Activation System." *Thromb. Res.* **27**: 289-301.
- Hartley, B.S. and Kauffman, D.L. (1966). "Corrections to the Amino Acid Sequence of Bovine Chymotrypsinogen A." *Biochem. J.* **101**: 229-231.
- Hartley, B.S. and Shotton, D.M. (1971). In: "The Enzymes" Vol 3, Boyer, R.D.(ed). Academic Press, New York. pp. 323-373.
- Hassam, H.J., Leonardi, A., Guerriero, R., Chelucci, C., Cianetti, L., Ciavarella, N., Ranieri, P., Pilolli, D., and Peschle, C. (1985). "Hemophilia B with Inhibitor: Molecular Analysis of the Subtotal Deletion of the Factor IX gene." *Blood* **66**: 728-730.
- Hay, C.W., Robertson, K.A., Yong, S.L., Thompson, A.R., Growe, G.H., and MacGillivray, R.T.A. (1986). "Use of a BamH1 polymorphism in the factor IX gene for the determination of hemophilia B carrier status." *Blood* **67** : 1508-1511.
- Hedner, U. and Davie, E.W. (1982). "Factor IX." in Colman, R.W., Hirsch, J., Marder, V.J., and Salzman, E.W. (eds): "Hemostasis and Thrombosis." J.B. Lippincott, Philadelphia.

- Hirsch, J. and Brain, E.A. (1979). "Hemostasis and Thrombosis: A Conceptual Approach." Churchill Livingstone, New York.
- Hougie, C. and Twomey, J.J. (1967). "A new type of factor IX deficiency." *Lancet* **1**: 698.
- Jackson, C.M. and Nemerson, Y. (1980). "Blood coagulation." *Ann. Rev. Biochem.* **49**: 765-811.
- James, M.N.G., Debaere, L.T.J., and Brayer, G.D. (1978). "Amino acid sequence alignment of bacterial and mammalian pancreatic serine proteases based on topological equivalences." *Can. J. Biochem.* **56**: 396-402.
- Kan, Y.N. and Dozy, A.M. (1978). "Polymorphism of DNA Sequence Adjacent to Human γ - Globin Structural Gene: Relationship to Sickle Mutation." *Proc. Natl. Acad. Sci. USA* **75**: 5631-5635.
- Kasper, C.K., Osterud, B., Minami, J.Y., Shonick, W., and Rapaport, S.I. (1977). "Hemophilia B: Characterization of Genetic Variants and Detection of Carriers." *J. Am. Soc. Hematol.* **50**: 351-366.
- Kraut, J. (1977). "Serine Proteases: Structure and Mechanism of Catalysis." *Ann. Rev. Biochem.* **46**: 331-358.
- Kurachi, K., and Davie, E.W. (1982). "Isolation and characterization of a cDNA coding for human factor IX." *Proc. Natl. Acad. Sci. USA* **79**: 6461-6464.
- Kurachi, K., Fujikawa, K., Schmer, G., and Davie, E.W. (1976). "Inhibition of Bovine Factor IX_a and Factor X_a by Antithrombin III" *Biochem.* **15**: 373-377.
- Lindquist, P.A., Fujikawa, K., and Davie, E.W. (1978). "Activation of Bovine Factor IX (Christmas Factor) by Factor XI_a (Activated Plasma Thromboplastin Antecedent) and a Protease from Russell's Viper Venom." *J. Biol. Chem.* **253**: 1902-1909.
- Link, R.P. and Castellino, F.J. (1983). "Kinetic Properties of Bovine Blood Coagulation Factors IX_α and IX_{αβ} Toward Synthetic Substrates." *Biochem.* **22**: 999-1005.
- Lollar, P.A. and Fass, D.N. (1984). "Inhibition of Activated Porcine Factor IX by Dansyl-glutamyl-glycyl-arginyl-chloromethylketone." *Arch. Biochem. Biophys.* **233**: 438-446.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y.

- Marquart, M., Walter, J., Deisenhofer, J., Bode, W., and Haber, R. (1983). "The Geometry of the Reactive Site and of the Peptide Groups in Trypsin, Trypsinogen, and Its Complexes with Inhibitors." *Acta Crystallogr. B.* **39**: 480.
- McGraw, R.A., Davis, L.M., Noyes, C.M., Lundblad, R.L., Roberts, H.R., Graham, J.B., and Stafford, D.W. (1985). "Evidence for a prevalent dimorphism in the activation peptide of human coagulation factor IX." *Proc. Natl. Acad. Sci. USA* **82**: 2847-2851.
- McKnight, S.L. and Kingsbury, R. (1982). "Transcriptional Control Signals of a Eukaryotic Protein-Coding Gene." *Science* **217**: 316-324.
- McMullen, B.A., Fujikawa, K., and Kisiel, W. (1978). "The occurrence of β -hydroxyaspartic acid in the vitamin K - dependent blood coagulation zymogens." *Biochem. Biophys. Res. Commun.* **115**: 8 -14.
- McRae, B.J., Kurachi, K., Heimark, R.L., Fujikawa, K., Davie, E.W., and Powers, J.C. (1981). "Mapping the Active Sites of Bovine Thrombin, Factor IX_a, Factor X_a, Factor XI_a, Factor XII_a, Plasma Kallikrein, and Trypsin with Amino Acid and Peptide Thioesters: Development of New Sensitive Substrates." *Biochem.* **20**: 7196-7206.
- Mertens, K., Cupers, R., van der Linden, I.K., and Bertina, R.M. (1983). "The Functional Defect of Factor IX Eindhoven, A Genetic Variant of Factor IX." *Thrombo. Hemostas.* **50**: 249 (Abstr. 783).
- Messing, J. (1983). "New M13 Vectors for Cloning." *Meth. Enzymol.* **101**: 20-78.
- Messing, J., Crea, R., and Seeburg, P.H. (1981). "A System for Shotgun DNA Sequencing." *Nucleic Acids Res.* **9**: 309-321.
- Morita, T., Isaacs, B.S., Esmon, C.T., and Johnson, A.E. (1984). "Derivatives of Blood Coagulation Factor IX Contain a High Affinity Ca^{2+} - Binding Site That Lacks γ -Carboxyglutamic Acid." *J. Biol. Chem.* **259**: 5698-5704.
- Mount, S.M. (1982). "A Catalogue of Splice Junction Sequences." *Nucleic Acids Res.* **10**: 459- 472.
- Noyes, C.M., Griffith, M.J., Roberts, H.R., and Lundblad, R.L. (1983). "Identification of the molecular defect in factor IX_{Chapel Hill}: Substitution of histidine for arginine at position 145." *Proc. Natl. Acad. Sci. USA* **80**: 4200-4202.
- Osterud, B., Kasper, C.K., Lavine, K.K., Prodanos, C., and Rapaport, S.I. (1981). "Purification and Properties of an Abnormal Blood Coagulation Factor IX (Factor IX_{Bm}) / Kinetics of its Inhibition of Factor X Activation by Factor VII and Bovine Tissue Factor." *Thrombos. Haemostas.* **45**: 55-59.

- Osterud, B., Kasper, C.K., and Prodanos, C. (1979). "Factor IX Variants of Hemophilia B. The Effect of Activated Factor XI and the Reaction Product of Factor VII and Tissue Factor on the Abnormal Factor IX Molecules." *Thromb. Res.* **15**: 235-243.
- Osterud, B.J., and Rapaport, S.I. (1977). "Activation of Factor IX by the reaction product of tissue factor and Factor VII: Additional pathway for initiating blood coagulation." *Proc. Natl. Acad. Sci. USA* **74**: 5260-5264.
- Peake, I.R., Furlong, B.L., and Bloom, A.L. (1984). "Carrier Detection by Direct Gene Analysis in a Family with Haemophilia B (Factor IX Deficiency)." *Lancet* **i**: 242-243.
- Proudfoot, N.J. and Brownlee, G.G. (1976). "3' Non-Coding Region Sequences in Eukaryotic Messenger RNA." *Nature* **263**: 211-214.
- Purrello, M., Alhadeff, B., Esposito, D., Szabo, P., Rocchi, M., Truett, M., Masiarz, F., and Siniscalco, M. (1985). "The human genes for hemophilia A and hemophilia B flank the X chromosome fragile site at Xq27.3" *EMBO J.* **4**: 725-729.
- Ratnoff, O.D. (1978). "Hereditary Disorders of Hemostasis." In: Stanbury, J.B., Wyngaarden, J.B., and Fredrickson, D.S. eds. "The Metabolic Basis of Inherited Disease." McGraw-Hill, New York.
- Rees, D.J.G., Rizza, C., and Brownlee, G.G. (1985). "Haemophilia B caused by a point mutation in a donor splice junction of the human factor IX gene." *Nature* **316**: 643-645.
- Roberts, H.R., Grizzle, J.E., McLester, W.D., and Penick, G.D. (1968). "Genetic variants of hemophilia B: Detection by means of a specific PTC inhibitor." *J. Clin. Invest.* **47**: 360.
- Rosenberg, J.S., McKenna, P.W. and Rosenberg, R.D. (1975). "Inhibition of Human Factor IX_a by Human Antithrombin." *J. Biol. Chem.* **250**: 8883-8888.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977). "DNA Sequencing with Chain-Terminating Inhibitors." *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.
- Sawyer, L., Shotton, D.M., Campbell, J.W., Wendell, P.L., Muirhead, H., Watson, H.C., Diamond, R., and Ladner, R.C. (1978). "The Atomic Structure of Crystalline Porcine Pancreatic Elastase at 2.5Å Resolution: Comparisons with the Structure of α -Chymotrypsin" *J. Mol. Biol.* **118**: 137-208.
- Sharp, P.A. (1981). "Speculations on RNA Splicing." *Cell* **23**: 643-646.
- Southern, E.M. (1975). "Detection of a Specific Sequence among DNA Fragments Separated by Gel Electrophoresis." *J. Mol. Biol.* **98**: 503-517.

- Steiner, D.F., Quinn, P.S., Chan, S.J., Marsh, J., and Tager, H.S. (1980). "Processing Mechanisms in the Biosynthesis of Proteins." *Ann. N. Y. Acad. Sci.* **343**: 1-16.
- Stormorken, H. (1979). "Advantages and Disadvantages of assaying coagulation and fibrinolytic parameters using chromogenic substrates." in Scully, M.F. and Kakkar, V.V. (eds): "Chromogenic Peptide Substrates: Chemistry and Clinical Usage." Churchill Livingstone, New York. pp. 32-58.
- Suomela, H., Blomback, M. and Blomback, B. (1977). "The Activation of Factor X Evaluated By Using Synthetic Substrates." *Thromb. Res.* **1**: 267-281.
- Suttie, J.W. (1985). "Vitamin K - Dependent Carboxylase." *Ann. Rev. Biochem.* **54**: 459-477.
- Thompson, A.R. (1983). "Monoclonal Antibody to an Epitope on the Heavy Chain of Factor IX Missing in Three Hemophilia-B Patients." *Blood* **62**: 1027-1034.
- Thompson, A.R. (1986). "Structure, Function, and Molecular Defects of Factor IX." *Blood* **67**: 565-572.
- Turner, A.D., Monroe, D.M., Roberts, H.R., Porter, N.A., and Pizzo, S.V. (1986). "p-Amidino Esters as Irreversible Inhibitors of Factor IX_a and X_a and Thrombin." *Biochem.* **25**: 4929-4935.
- Usharani, P., Warn-Cramer, J., Kasper, C.K., and Bajaj, S.P. (1985). "Characterization of Three Abnormal Factor IX Variants (B_M Lake Elsinore, Long Beach, and Los Angeles) of Hemophilia-B: Evidence for Defects Affecting the Catalytic Site." *J. Clin. Invest.* **75**: 76-83.
- Vidaud, M., Chabret, C., Gazengel, C., Grunebaum, L., Cazenave, J.P., and Goossens, M. (1986). "A De Novo Intragenic Deletion of the Potential EGF Domain of the Factor IX Gene in a Family With Severe Hemophilia B." *Blood* **68**: 961-963.
- von Heijne, G. (1983). "Patterns of Amino Acids Near Signal-Sequence Cleavage Sites." *Eur. J. Biochem.* **133**: 17-21.
- von Heijne, G. (1985). "Signal Sequences: The Limits of Variation." *J. Mol. Biol.* **184**: 99-105.
- Wallmark, A., Ljung, R., Nilsson, I.M., Holmberg, L., Hedner, U., Lindvall, M., and Sjogren, H-O. (1985). "Polymorphism of normal factor IX detected by mouse monoclonal antibodies." *Proc. Natl. Acad. Sci. USA* **82**: 3839-3843.
- Weatherall, D.J. (1985). "The new genetics and clinical practice." 2nd edn. Oxford University Press.
- Weatherall, D.J., Higgs, D.H., Wood, W.G., and Clegg, J.B. (1984) "Genetic disorders of human haemoglobin as models for analysing gene regulation." *Phil. Trans. R. Soc.* **B307**: 247-259.

- Winship, P.R., Anson, D.S., Rizza C.R., Brownlee, G.G. (1984). "Carrier detection in haemophilia B using two further intragenic restriction fragment length polymorphisms." *Nucleic Acids Res.* **12**: 8861-8872.
- Winship, P.R. and Brownlee, G.G. (1986). "Diagnosis of Haemophilia B Carriers Using Intragenic Oligonucleotide Probes." *Lancet* **II**: 218.
- Yoshioka, A., Ohkubo, Y., Nishimura, T., Tanaka, I., Fukui, H., Ogata, K., Kamiya, T., and Takahashi, H. (1986). "Heterogeneity of Factor IX BM: Difference of Cleavage Sites by Factor XI_a and Ca²⁺ in Factor IX Kashihara, Factor IX Nagoya and Factor IX Niigata." *Thrombos. Res.* **42**: 595-604.
- Yoshitake, S., Schach, B.G., Foster, D.C., Davie, E.W., and Kurachi, K. (1985). "Nucleotide Sequence of the Gene for Human Factor IX (Antihemophilic Factor B)". *Biochemistry* **24**: 3736-3750.