Volume 11 Number 8 1983 Nucleic Acids Research

Isolation of a human anti-haemophilic factor IX cDNA clone using a unique 52-base synthetic oligonucleotide probe deduced from the amino acid sequence of bovine factor IX

Michael Jaye, Henri de la Salle, Fabienne Schamber, Alain Balland, Vipin Kohli, Annie Findeli, Paul Tolstoshev and Jean-Pierre Lecocq.

Transgène S.A., 11 rue Humann, 67000 Strasbourg, France

Received 22 February 1983; Accepted 22 March 1983

ABSTRACT

A unique 52mer oligonucleotide deduced from the amino acid sequence of bovine Factor IX was synthesized and used as a probe to screen a human liver cDNA bank. The Factor IX clone isolated shows 5 differences in nucleotide and deduced amino acid sequence as compared to a previously isolated clone. In addition, precisely one codon has been deleted.

INTRODUCTION

Factor IX (Christmas Factor) is a coagulation factor of the intrinsic pathway of blood clotting. It is synthesized as a zymogen, and is modified post-translationally both by the addition of carbohydrate and by the conversion of 12 residues of glutamic acid to γ-carboxyglutamic acid, the latter modification being a vitamin K-dependent process (1). By analogy with other coagulation factors and serum proteins, the site of synthesis of factor IX is assumed to be the liver. Factor IX deficiency (Hemophilia B or Christmas Disease) is manifested by dangerously prolonged bleeding times, and is transmitted as a sex-linked recessive trait (2). The molecular nature of Factor IX deficiency is unknown.

We wished to obtain a cDNA clone of Human Factor IX, while at the same time attempting a novel method of screening cDNA banks for specific sequences. Based on the known amino-acid sequence of bovine Factor IX, we attempted to design a unique oligonucleotide to be used as a probe for the corresponding cDNA clone. In designing this unique oligonucleotide, the following variables were considered : 1) the segment of the Factor IX amino acid sequence with the least ambiguity in the corresponding codons, 2) codon usage in already sequenced bovine cDNAs corresponding to proteins synthesized and secreted by the liver, 3) the relative stability of G:T versus G:A mismatches while screening the cDNA bank with the oligonucleotide by colony hybridization, and 4) the sequence of the oligonucleotide probe permitting

minimum predictable secondary structure in the oligonucleotide. This strategy differs from other strategies currently employed in our laboratory to screen cDNA banks for unique sequences since one long oligonucleotide is used as specific probe, rather than mixtures of shorter oligonucleotides which cover all possible arrangements of codons. The latter strategy has recently been used to obtain Factor IX cDNA clones by Choo et al. (3) and Kurachi and Davie (4).

MATERIALS AND METHODS

Isolation of liver poly-A-containing RNA and cDNA synthesis.

Livers were obtained post mortem and quickly frozen in liquid nitrogen. RNA was prepared from 5 gram of liver using the 8 M guanidium hydrochloride extraction procedure, essentially as described by Tolstoshev et al. (5). The RNA thus obtained was chromatographed on a polyU-Sepharose (Pharmacia) column, as described by the manufacturer to enrich for poly-A-containing RNA which served as a template for reverse transcription using oligo(dT) as a primer. cDNA was synthesized in a 100 µl reaction containing 100 mM Tris-HCl, pH 8.3, 10 mM $MgCl_2$, 50 mM KCl, 30 mM β-mercaptoethanol, 10 ug/ml $oligo(dT)_{12-18}$, 50 ug/ml poly-A-RNA, 0.5 mM each of dATP, dGTP, dTTP, and dCTP, 80 units of avian myoblastosis virus reverse transcriptase (Life Sciences Inc., St Petersberg, Florida). After 45 minutes at 42°C, the reaction was terminated and cDNA-template complexes simultaneously denatured by heating at 105°C for 3 minutes followed by rapid transfer to an ice bath.

For the synthesis of the second DNA strand, the above reaction was diluted 5-fold and adjusted to final concentrations of 100 mM HEPES-KOH, pH 6.9, 100 mM KCl, 200 uM each of dATP, dGTP, dTTP, and ^{32}P-α-dCTP (specific activity 0.5 Ci/mmole). 10 units of E. coli DNA polymerase (Klenow fragment, Boehringer Mannheim) was then added, and the incubation held at 25°C for 2 hours.

Double stranded cDNA (dscDNA) was extracted with an equal volume of phenol : chloroform (50:50) saturated with 10 mM Tris-HCl pH 8.0, 1 mM EDTA, and precipitated with ethanol. Approximately 970 ng of dscDNA was obtained from 5 µg poly-A-RNA. This was rendered blunt-ended by digestion with 5 units of S1 nuclease in a reaction volume of 0.1 ml containing 30 mM sodium acetate, pH 4.8, 300 mM NaCl, 3 mM $ZnCl_2$. After 1 hour at 37°C, EDTA and SDS were added to final concentrations of 10 mM and 0.1 %, respectively and the reaction heated at 65°C for 5 minutes. The S1-digested dscDNA was then

applied to preformed 5-20 % sucrose gradient containing 100 mM Tris-HCl pH 7.5, 5 mM EDTA, 100 mM NaCl, and centrifuged at 30,000 rpm at 15°C for 16 hours in a SW60 Ti rotor. Fractions of 0.5 ml were collected. To determine the size of the cDNA, 1 µl of each fraction was electrophoresed on neutral agarose gels and the migration of each fraction compared with appropriate molecular weight markers. Fractions containing cDNA larger than 1 kilobase were pooled and precipitated with 2 volumes of ethanol overnight in the presence of 5 µg of E. coli tRNA as carrier. The pellet was resuspended in 20 µl 0.1 x SSC (15 mM NaCl, 0.15 mM sodium citrate pH 7.0).

cDNA cloning.

Homopolymeric tails (approximately 15 dCMP residues) were added to the 3' ends of the cDNA using terminal deoxynucleotidyl transferase (6). Similarly, homopolymeric dGMP tails (approximately 13 dGMP residues) were added to the 3' extremities of plasmid pBR322 previously digested with the restriction nuclease Pst-I.

60 ng of dCMP tailed cDNA and 1 µg of dCMP tailed vector were annealed for 2 hours at 42°C in 10 mM Tris-HCl, pH 7.5, 100 mM NaCl. Recombinant plasmids were then transformed into E. coli strain 8739 (7), and transformants selected on LB-agar plates containing 15 µg/ml tetracycline. 30,000 transformants were obtained, of which approximately 50 % contained cDNA inserts larger than 1 kilobase. All transformants were scraped from the plates into a minimal volume of LB, and after addition of an equal volume of glycerol, the "bank" was stored at - 20°C.

Selection and synthesis of oligonucleotide probe for Factor IX.

A total of 1755 nucleotides of coding sequence corresponding to bovine proteins synthesized by the liver (8, 9) were analyzed with regard to codon useage as described below and a sequence of 52 bases was chosen. This oligonucleotide was constructed as follows :

Oligonucleotides A d(CTCACACTGATCACCATCCACATACT),

B d(GCTTCCAGAACTCTGTAGTCTTCTCA),

and the complementary fragment C d(GGAAGCAGTATGT) were chemically synthesized by the phosphotriester method on an inorganic solid support as described previously (10) and purified by HPLC (11). 0.5 nmole oligonucleotide B was incubated for 30 minutes at 37°C in a 10 µl reaction containing 60 mM Tris-HCl pH 7.8, 6 mM $MgCl_2$, 6 mM dithiothreitol, 0.1 mm ATP, 6 pmole ^{32}P-γ-ATP (3000 Ci/mmole), 2 units T4 polynucleotide kinase. The 5' phosphorylated oligomer B was mixed with 0.5 nmole oligonucleotides A

and C in 10 mM Tris HCl pH 7.5, 1 mM EDTA, heated to 100°C and the mixture annealed by slow cooling to 4°C. The annealed oligonucleotides A and B were ligated in a 50 µl reaction containing 66 mM Tris-HCl pH 7.5, 100 mM NaCl, 7.5 mM $MgCl_2$, 2 mM dithiothreitol, 0.5 mM spermidine, 0.2 mM EDTA, 4 units T4 DNA ligase at 4°C overnight. The resulting 52-mer was then purified from the ligation mixture by electrophoresis on a 20 % polyacrylamide gel.

Screening of liver cDNA banks.

Approximately 10,000 colonies were grown overnight on LB agar plates containing 15 µg/ml tetracycline. The next day, the bacterial colonies were transferred to nitrocellulose filters and the filters prepared for colony hyridization essentially as described (12). The bacteria remaining on the original tetracycline plates were regrown for a few hours at 37°C.

The hybridization probe was prepared by kination of 100 ng 52-mer with 50 uCi of ^{32}P-γ-ATP. Hybridization was performed at 48°C overnight in 40 ml of 6 x SSC, 1 x Denhardt's solution, 0.1 % SDS, 50 µg/ml E. coli tRNA. The next day, filters were washed at 48°C with several changes of 6 x SSC, 0.1 % SDS. Filters were then dried, and autoradiographed overnight.

RESULTS

Design of the oligonucleotide probe.

The amino acid sequence of bovine Factor IX (13) was analyzed to find the longest region corresponding to the least codon ambiguity. We reasoned that if regions of amino acid sequence containing amino acids encoded by 6 codons were avoided in designing the oligonucleotide probe, the degeneracy of codons would be limited to the third position and thus the probe would be at least 66 % homologous. The probe was thus designed to correspond to amino acids 35 to 52 of bovine Factor IX, which consists of 5 amino acids encoded by 4 codons (thr, val, gly), 12 amino acids encoded by 2 codons (gln, lys, phe, gln, tyr, asp, cys), and one amino acid encoded by 1 codon (trp) (Fig. 1).

Analysis of 585 codons of bovine prothrombin and fibrinogen revealed that the valine codon GTG is used in 21/38 instances. Similarly, TTC and TGT are used 2 times more frequently than the other degenerate codon in the case of amino acids phe and cys, respectively. No other clear preferences were obvious, however certain nucleotides were infrequently seen in the third position of certain codons, for example G in the case of both glycine and threonine.

Next, we considered that a G:T mismatch, although less stable than a G:C

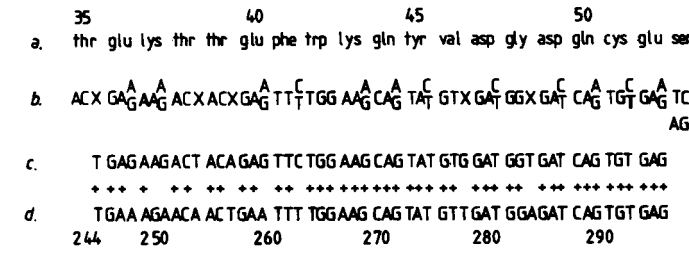

Figure 1 : Design of the Unique Oligonucleotide Probe
a. amino acid sequence of bovine Factor IX residues 35-53 (reference 13).
b. degenerate nucleotide sequence deduced from a.
c. selected sequence of oligonucleotide probe from b.
d. sequence of nucleotides 244-295 of pTG397 corresponding to amino acids 35-53 of the mature form of Factor IX (see figure 4).
+ indicates homology between nucleotide sequence of the oligonucleotide probe and the cDNA sequence.

basepair, would at least contribute to the stability of the hybridization of the 52mer during colony hybridization, whereas an A:C mismatch would not. Thus, G was chosen for each amino acid (glu, lys, gln) for which the choice in the third position of the degenerate codon was either G or A; similarly, T was chosen over A in the case of amino acids tyr and asp. For these considerations, we confined our choice to T or G as nucleotide in the third position for threonine and glycine, however, as mentioned above, G was infrequently seen in the third position of threonine codons. Thus, we chose T for one threonine codon, and, to avoid possible problems of secondary structure within the 52mer, chose T in the third position for glycine and A in the third position of a second threonine codon.

Screening of liver cDNA banks and identification of a human Factor IX clone.

At the time that our 52mer probe was designed, our original strategy was to utilize the 52mer to screen a bovine liver cDNA bank and then to employ the bovine Factor IX cDNA clone as probe in a human liver cDNA bank. This strategy was based on the known homology in amino acid sequence at the NH_2-terminus of the bovine and human proteins (14).

Relying upon this interspecies homology in amino acid sequence, we screened, in parallel, size-fractionated bovine and human liver cDNA banks. Relatively low stringency conditions of hybridization and washing were employed, since the actual homology between the 52mer probe and Factor IX cDNA was unknown. Several colonies in the bovine cDNA bank gave strong positive signals when screened with the 52-mer probe. In the human liver cDNA bank,

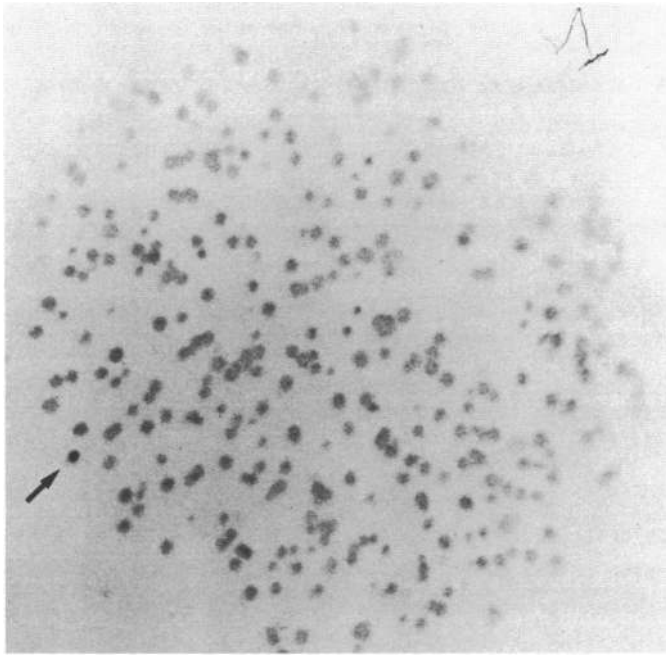

Figure 2 : Autoradiograph of Human Liver cDNA Bank Screened with Kinated 52-mer Probe. Hybridization and washing conditions are described in Materials and Methods. The arrow indicates the Factor IX cDNA clone obtained.

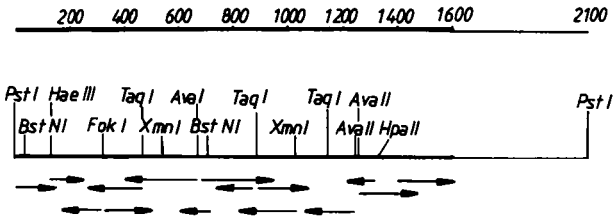

Figure 3 : Restriction Enzyme Map of cDNA Insert of pTG397, and Sequencing Strategy. Nucleotide sequences deriving from the PstI site at position 0, the XmnI site at position 1029, and the HpaII site at position 1334 were determined using dideoxynucleotides as chain terminators (ref. 16) after subcloning of the appropriate restriction fragments into bacteriophages M13mp8 or M13mp9 (ref. 15).
All other sequence was performed as described by Maxam and Gilbert (17).
Only the HaeIII site actually used in the sequencing strategy is indicated.
Nucleotides 1640-2100, which have not been sequenced, are indicated by a thin line.

one colony was obtained which gave a strong signal under the conditions employed (Fig. 2). Since our objective was to obtain a human Factor IX cDNA clone, no further analysis was performed upon the bovine clones and we focused our effort on the human clone.

To establish whether the colony selected from the human liver cDNA bank corresponded to a Factor IX clone, the recombinant plasmid was digested with PstI and the excised cDNA insert transferred to M13mp8 (15) and sequenced using dideoxynucleotides as chain terminators (16). The sequence obtained corresponded to nucleotides 1-180 of figure 3. The deduced amino acid sequence confirmed that the clone was a cDNA clone of human Factor IX. Further sequencing of the cDNA insert of this clone (pTG397) was performed both as described by Maxam and Gilbert (17) and by subcloning of several restriction fragments into M13mp9 (15). The sequencing strategy is presented in figure 3. The sequence obtained is shown in figure 4.

DISCUSSION

Employing a unique oligonucleotide probe, we have isolated a clone for human Factor IX containing an insert of 2.1 kb. When the 2.1 kb insert is used as a probe to rescreen the human liver cDNA bank, another Factor IX specific clone (pTG398), containing an insert of 2.6 kb, was isolated (data not shown). The 5' end of pTG398 is located at nucleotide 480 in figure 4. This suggests that the 3' non-translated end of pTG397 is incomplete, and that the complete 3'non-translated end has a length of approximately 1.7 kb.

The sequence of the entire coding region and 250 nucleotides of 3' non-coding region of pTG397 are presented in figure 4. Greater than 75 % of this sequence has been established on both strands. We observe several differences between our nucleotide sequence and that found by Kurachi and Davie (4). These differences involve transitions at nucleotides 38 (A→G), 581 (G→A), 611 (C→T), and 650 (G→A), one inversion at nucleotides 249-250 (AG→GA), and a deletion of an alanine codon (GCA) following position 895 in our sequence (Fig. 5). All but one of these differences (the first transition) have been confirmed on both strands. The first transition generates a Bst NI restriction site, which has been confirmed (data not shown). None of these differences are silent. They result in one change in the signal peptide (serine to glycine), two amino acid changes in the eventual activated Factor IX molecule (lysine to arginine, and a deletion of an alanine residue), and 3 amino acid changes (alanine to threonine, proline to serine, and glycine to serine) in the activation peptide region. These data suggest

```
1                         25                                     50                                  75
T ATG CAG CGC GTG AAC ATG ATC ATG GCA GAA TCA CCA GGC CTC ATC ACC ATC TGC CTT TTA GGA TAT CTA CTC AGT
  met gln arg val asn met ile met ala glu ser pro gly leu ile thr ile cys leu leu gly tyr leu leu ser

                          100                                    125                                 150
  GCT GAA TGT ACA GTT TTT CTT GAT CAT GAA AAC GCC AAC AAA ATT CTG AAT CGG CCA AAG AGG TAT AAT TCA GGT
  ala glu cys thr val phe leu asp his glu asn ala asn lys ile leu asn arg pro lys arg tyr asn ser gly

                          175                                    200                                 225
  AAA TTG GAA GAG TTT GTT CAA GGG AAC CTT GAG AGA GAA TGT ATG GAA GAA AAG TGT AGT TTT GAA GAA GCA CGA
  lys leu glu glu phe val gln gly asn leu glu arg glu cys met glu glu lys cys ser phe glu glu ala arg

                          250                                    275                                 300
  GAA GTT TTT GAA AAC ACT GAA AGA ACA ACT GAA TTT TGG AAG CAG TAT GTT GAT GGA GAT CAG TGT GAG TCC AAT
  glu val phe glu asn thr glu arg thr thr glu phe trp lys gln tyr val asp gly asp gln cys glu ser asn

                          325                                    350                                 375
  CCA TGT TTA AAT GGC GGC AGT TGC AAG GAT GAC ATT AAT TCC TAT GAA TGT TGG TGT CCC TTT GGA TTT GAA GGA
  pro cys leu asn gly gly ser cys lys asp asp ile asn ser tyr glu cys trp cys pro phe gly phe glu gly

                          400                                    425                                 450
  AAG AAC TGT GAA TTA GAT GTA ACA TGT AAC ATT AAG AAT GGC AGA TGC GAG CAG TTT TGT AAA AAT AGT GCT GAT
  lys asn cys glu leu asp val thr cys asn ile lys asn gly arg cys glu gln phe cys lys asn ser ala asp

                          475                                    500                                 525
  AAC AAG GTG GTT TGC TCC TGT ACT GAG GGA TAT CGA CTT GCA GAA AAC CAG AAG TCC TGT GAA CCA GCA GTG CCA
  asn lys val val cys ser cys thr glu gly tyr arg leu ala glu asn gln lys ser cys glu pro ala val pro

                          550                                    575     *                           600
  TTT CCA TGT GGA AGA GTT TCT GTT TCA CAA ACT TCT AAG CTC ACC CGT GCT GAG ACT GTT TTT CCT GAT GTG GAC
  phe pro cys gly arg val ser val ser gln thr ser lys leu thr arg ala glu thr val phe pro asp val asp

              *           625                                    650                                 675
                                                                     *
  TAT GTA AAT TCT ACT GAA GCT GAA ACC ATT TTG GAT AAC ATC ACT CAA AGC ACC CAA TCA TTT AAT GAC TTC ACT
  tyr val asn ser thr glu ala glu thr ile leu asp asn ile thr gln ser thr gln ser phe asn asp phe thr

                          700                                    725                                 750
  CGG GTT GTT GGT GGA GAA GAT GCC AAA CCA GGT CAA TTC CCT TGG CAG GTT GTT TTG AAT GGT AAA GTT GAT GCA
  arg val val gly gly glu asp ala lys pro gly gln phe pro trp gln val val leu asn gly lys val asp ala

                          775                                    800                                 825
  TTC TGT GGA GGC TCT ATC GTT AAT GAA AAA TGG ATT GTA ACT GCT GCC CAC TGT GTT GAA ACT GGT GTT AAA ATT
  phe cys gly gly ser ile val asn glu lys trp ile val thr ala ala his cys val glu thr gly val lys ile

                          850                                    875                           Δ   900
  ACA GTT GTC GCA GGT GAA CAT AAT ATT GAG GAG ACA GAA CAT ACA GAG CAA AAG CGA AAT GTG ATT CGA ATT ATT
  thr val val ala gly glu his asn ile glu glu thr glu his thr glu gln lys arg asn val ile arg ile ile

                          925                                    950                                 975
  CCT CAC CAC AAC TAC AAT GCA GCT ATT AAT AAG TAC AAC CAT GAC ATT GCC CTT CTG GAA CTG GAC GAA CCC TTA
  pro his his asn tyr asn ala ala ile asn lys tyr asn his asp ile ala leu leu glu leu asp glu pro leu

                          1000                                   1025                                1050
  GTG CTA AAC AGC TAC GTT ACA CCT ATT TGC ATT GCT GAC AAG GAA TAC ACG AAC ATC TTC CTC AAA TTT GGA TCT
  val leu asn ser tyr val thr pro ile cys ile ala asp lys glu tyr thr asn ile phe leu lys phe gly ser

                          1075                                   1100                                1125
  GGC TAT GTA AGT GGC TGG GGA AGA GTC TTC CAC AAA GGG AGA TCA GCT TTA GTT CTT CAG TAC CTT AGA GTT CCA
  gly tyr val ser gly trp gly arg val phe his lys gly arg ser ala leu val leu gln tyr leu arg val pro

                          1150                                   1175                                1200
  CTT GTT GAC CGA GCC ACA TGT CTT CGA TCT ACA AAG TTC ACC ATC TAT AAC AAC ATG TTC TGT GCT GGC TTC CAT
  leu val asp arg ala thr cys leu arg ser thr lys phe thr ile tyr asn asn met phe cys ala gly phe his

                          1225                                   1250                                1275
  GAA GGA GGT AGA GAT TCA TGT CAA GGA GAT AGT GGG GGA CCC CAT GTT ACT GAA GTG GAA GGG ACC AGT TTC TTA
  glu gly gly arg asp ser cys gln gly asp ser gly gly pro his val thr glu val glu gly thr ser phe leu

                          1300                                   1325                                1350
  ACT GGA ATT ATT AGC TGG GGT GAA GAG TGT GCA ATG AAA GGC AAA TAT GGA ATA TAT ACC AAG GTA TCC CGG TAT
  thr gly ile ile ser trp gly glu glu cys ala met lys gly lys tyr gly ile tyr thr lys val ser arg tyr

                          1375                                   1400                                1425
  GTC AAC TGG ATT AAG GAA AAA ACA AAG CTC ACT TAA TGA AAG ATG GAT TTC CAA GGT TAA TTC ATT GGA ATT GAA
  val asn trp ile lys glu lys thr lys leu thr --- ---

                          1450                                   1475                                1500
  AAT TAA CAG GGC CTC TCA CTA ACT AAT CAC TTT CCC ATC TTT TGT TAG ATT TGA ATA TAT ACA TTC TAT GAT CAT

                          1525                                   1550                                1575
  TGC TTT TTC TCT TTA CAG GGG AGA ATT TCA TAT TTT ACC TGA GCA AAT TGA TTA GAA AAT GGA ACC ACT AGA GGA

                          1600                                   1625
  ATA TAA TGT GTT AGG AAA TTA CAG TCA TTT CTA AGG GCC CAG CCT TGA CAA ATT GTG AGT AAA ...............
```

Figure 4 : Nucleotide and Deduced Amino Acid Sequence of cDNA Insert of pTG397. ... represents nucleotides 1640-2100, which have not been sequenced. Asterisks indicate the positions of divergence between the nucleotide sequence of this clone and that previously published (reference 4). Δ represents the position of the missing alanine codon (see Discussion).

that, like other serum proteins (18) Factor IX is polymorphic. The G→A transition at nucleotide 581, from which we deduce threonine, rather than alanine in the connecting peptide, corresponds to the presence of threonine in this position, as determined by amino acid sequence (19). It is interest

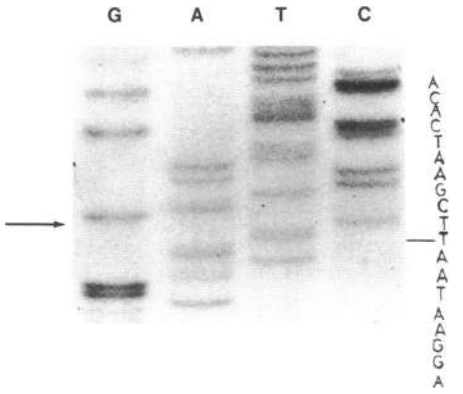

Figure 5 : Nucleotide Sequence around the Missing Alanine Codon
The sequence read 5'...AGGAATAATTCGAATCACA...3' corresponds to the reverse, complementary sequence of nucleotides 886-904 of figure 4.
The presence of the alanine codon GCA (reference 4) would cause this sequence to be read 5'...AGGAATAATTGCTCGAATCACA...3'

ing that the greatest divergence in nucleotide and protein sequence is found in the activation peptide. This may be due to a lesser degree of selective pressure on this portion of the gene, which corresponds to a peptide which is removed from Factor IX upon activation.

The deletion of an alanine codon following nucleotide 895 in our sequence is reminiscent of the rat prolactin gene, wherein the deletion of one codon in one of two cDNA sequences can be explained by differential splicing (20). Differential splicing probably does not explain the deletion of the alanine codon in our sequence, since no differential splice sites which follow the GU.....AG rules proposed by Breathnach et al. (21) are apparent in the sequence containing the extra codon.

Consideration of codon use, G:T mismatch, and elimination of possible secondary structure within the oligonucleotide probe has enabled the design of a unique 52mer which has been successfully used as probe for a specific cloned sequence. The 52mer which was deduced from the bovine sequence is homologous with the cloned human sequence in 43/52 nucleotides, while 2 mismatches are G:T basepairs. This precisely reflects the 85 % homology in interspecies nucleotide sequence observed previously (3). We have not rigorously examined signal strength or specificity during colony hybridization as a function of hybridization and washing conditions. These conditions

would inevitably vary depending upon the degree of homology between probe and cloned sequence as well as actual position and nature of the mismatched bases. The general approach of utilizing a long unique oligonucleotide, rather than a mixture of shorter oligonucleotides, should find broad application in probing for unique sequences.

ACKNOWLEDGEMENTS

The authors thank Richard Lathe for suggesting the use of a long synthetic oligonucleotide probe. We are grateful to Professors Pierre Chambon and Philippe Kourilsky for their continued interest in this work. Drs. J.P. Cazenave, T. Edgington, D. Fair, D. Meyer and J. Cinqualbre are thanked for many helpful discussions. The expert secretarial assistance of Patricia Gabriel is gratefully acknowledged.
This project is supported by the Institut Mérieux, Lyon.

REFERENCES

1. Bertina, R.M. and Veltkamp, J.J. (1981). in Hemostasis and Thrombosis, Bloom, A.L. and Thomas, D.T., Eds. pp. 98-110, Churchill Livingstone, London.
2. Hedner, U. and Davie, E.W. (1982). in Hemostasis and Thrombosis, Colman, R.W., Hirsh, J., Marder, V.J. and Salzman, E.W., Eds. pp. 29-38, J.B. Lippincott Co., Philadelphia.
3. Choo, K.H., Gould, K.G., Rees, D.J.G. and Brownlee, G.G. (1982). Nature 299, 178-180.
4. Kurachi, K. and Davie, E.W. (1982). Proc. Natl. Acad. Sci. U.S.A. 79, 6461-6464.
5. Tolstoshev, P., Berg, R.A., Rennard, S.I., Bradley, K.N., Trapnel, B.C. and Crystal, R.G. (1981). J. Biol. Chem. 256, 3135-3140.
6. Deng, G. and Wu, R. (1981). Nucl. Acids Res. 9, 4173-4188.
7. Murray, N.E., Brammar, W.J. and Murray, K. (1977). Mol. Gen. Genetics 150, 53-61.
8. MacGillivray, R.T.A, Degen, S.J.F., Chandra, T., Woo, S.L.C. and Davie, E.W. (1980). Proc. Natl. Acad. Sci. U.S.A. 77, 5153-5157.
9. Chung, D.W., Rixon, M.W., MacGillivray, R.T.A. and Davie, E.W. (1981). Proc. Natl. Acad. Sci. U.S.A. 78, 1466-1470.
10. Kohli, V., Balland, A., Wintzerith, M., Sauerwald, R., Staub, A. and Lecocq, J.P. (1982). Nucl. Acids Res. 10, 7439-7448.
11. Fritz, H.J., Belagaje, R., Brown, E.L., Fritz, R.H., Jones, R.A., Lees, R.F. and Khorana, H.G. (1978). Biochemistry 17, 1257-1267.
12. Grunstein, M. and Wallis, J. (1979). Methods in Enzymology 68, 379-389.
13. Katayama, K., Ericsson, L.H., Enfield, D.L., Walsh, K.A., Neurath, H., Davie, E.W. and Titani, K. (1979). Proc. Natl. Acad. Sci. U.S.A. 76, 4990-4994.
14. Di Scipio, R.G., Hermodson, M.A., Yates, S.G. and Davie, E.W. (1977). Biochemistry 16, 698-706.
15. Messing, J. and Vieira, J. (1982). Gene 19, 269-276.
16. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.

17.Maxam, A.M. and Gilbert, W. (1980). Methods in Enzymology 65, 499-555.
18.Cooper, D.W. (1978) in The Biochemical Genetics of Man, Brock, D.J.H. and Mayo, O., pp. 271-324, Eds. Academic Press, London.
19.Di Scipio, R.G., Kurachi, K. and Davie, E.W. (1978). J. Clin. Inves. 61, 1528-1538.
20.Maurer, R.A., Erwin, C.R. and Donelson, J.E. (1981). J. Biol. Chem. 256, 10524-10528.
21.Breathnach, R.A., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978). Proc. Natl. Acad. Sci. U.S.A. 75, 4853-4857.