Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., & Surrey, S. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4298–4302.
Proudfoot, N. J., & Brownlee, G. G. (1976) *Nature (London) 263*, 211–214.
Ratnoff, O. D. (1966) *Prog. Hematol. 5*, 204–245.
Revak, D. D., Cochrane, C. G., Johnson, A. R., & Hugli, T. E. (1974) *J. Clin. Invest. 54*, 619–627.
Revak, S. D., Cochrane, C. G., & Griffin, J. H. (1977) *J. Clin. Invest. 59*, 1167–1175.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schapira, M., Despland, E., Scott, C. F., Boxer, L. A., & Colman, R. W. (1982) *J. Clin. Invest. 69*, 1199–1202.
Scott, C. F., Silver, L. D., Schapira, M., & Colman, R. W. (1984) *J. Clin. Invest. 73*, 954–962.
Sealey, J. E., Atlas, S. A., Laragh, J. H., Silvergerg, M., & Kaplan, A. P. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5914–5918.
Shine, J., Mason, A. J., Evans, B. A., & Richards, R. I. (1983) *Cold Spring Harbor Symp. Quant. Biol. 48*, 419–426.
Sottrup-Jensen, L., Claeys, H., Zajdel, M., Petersen, T. E., & Magnusson, S. (1978) *Prog. Chem. Fibrinolysis Thrombolysis 3*, 191–209.
Takagaki, Y., Kitamura, N., & Nakanishi, S. (1985) *J. Biol. Chem. 260*, 8601–8609.
Tans, G., & Griffin, J. H. (1982) *Blood 59*, 69–75.
Titani, K., Hermodson, M. A., Fujikawa, K., Ericsson, L. H., Walsh, K. A., Neurath, H., & Davie, E. W. (1972) *Biochemistry 11*, 4899–4903.
van der Graaf, F., Tans, G., Bouma, B. N., & Griffin, J. H. (1982) *J. Biol. Chem. 257*, 14300–14305.
Weber, K., & Osborn, M. (1969) *J. Biol. Chem. 244*, 4406–4412.
Wilson, A. C., Carlson, S. S., & White, T. J. (1977) *Annu. Rev. Biochem. 46*, 573–639.
Yoshitake, S., Schach, B. G., Foster, D. C., & Davie, E. W. (1985) *Biochemistry 24*, 3736–3750.
Young, R. A., & Davis, R. W. (1983a) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1194–1198.
Young, R. A., & Davis, R. W. (1983b) *Science (Washington, D.C.) 222*, 778–782.

Amino Acid Sequence of Human Factor XI, a Blood Coagulation Factor with Four Tandem Repeats That Are Highly Homologous with Plasma Prekallikrein†

Kazuo Fujikawa, Dominic W. Chung, Lee E. Hendrickson, and Earl W. Davie*
Department of Biochemistry, University of Washington, Seattle, Washington 98195
Received November 21, 1985; Revised Manuscript Received January 10, 1986

ABSTRACT: A λgt11 cDNA library prepared from human liver poly(A) RNA has been screened with affinity-purified antibody to human factor XI, a blood coagulation factor composed of two identical polypeptide chains linked by a disulfide bond(s). A cDNA insert coding for factor XI was isolated and shown to contain 2097 nucleotides, including 54 nucleotides coding for a leader peptide of 18 amino acids and 1821 nucleotides coding for 607 amino acids that are present in each of the 2 chains of the mature protein. The cDNA for factor XI also contained a stop codon (TGA), a potential polyadenylation or processing sequence (AACAAA), and a poly(A) tail at the 3′ end. Five potential N-glycosylation sites were found in each of the two chains of factor XI. The cleavage site for the activation of factor XI by factor XII_a was identified as an internal peptide bond between Arg-369 and Ile-370 in each polypeptide chain. This was based upon the amino acid sequence predicted by the cDNA and the amino acid sequence previously reported for the amino-terminal portion of the light chain of factor XI. Each heavy chain of factor XI_a (369 amino acids) was found to contain 4 tandem repeats of 90 (or 91) amino acids plus a short connecting peptide. Each repeat probably forms a separate domain containing three internal disulfide bonds. The light chains of factor XI_a (each 238 amino acids) contain the catalytic portion of the enzyme with sequences that are typical of the trypsin family of serine proteases. The amino acid sequence of factor XI shows 58% identity with human plasma prekallikrein.

Factor XI (plasma thromboplastin antecedent) is a zymogen to a serine protease that participates in the early or contact phase of blood coagulation (Davie et al., 1979). Patients deficient in this plasma glycoprotein have a mild bleeding disorder or are asymptomatic (Rapaport et al., 1961; Ratnoff & Saito, 1979). The physical and chemical properties of factor XI have been determined with purified preparations isolated from human (Bouma & Griffin, 1977; Kurachi & Davie, 1977, 1981) and bovine plasma (Koide et al., 1976, 1977a; Kurachi et al., 1980). Factor XI is unique in that it is a zymogen to a serine protease that is composed of two identical polypeptide chains linked by a disulfide bond(s). Its molecular weight has been reported to be between 125 000 and 160 000 for the dimer and between 55 000 and 63 000 for the monomer. Human factor XI contains 5% carbohydrate (Kurachi & Davie, 1977), whereas bovine factor XI contains 11% carbohydrate (Koide et al., 1977a).

Factor XI circulates in plasma as a complex with high molecular weight kininogen (HMW kininogen)[1] (Thompson et al., 1977). It is converted in vitro to factor XI_a by factor

†This work was supported by Research Grant HL16919 from the National Institutes of Health. D.W.C. is an Established Investigator of the American Heart Association.

[1] Abbreviations: HMW kininogen, high molecular weight kininogen; SDS, sodium dodecyl sulfate; TBS, 50 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; kb, kilobase(s).

0006-2960/86/0425-2417$01.50/0

XII_a (or factor XII) in the presence of an anionic surface and HMW kininogen (Griffin & Cochrane, 1976; Bouma & Griffin, 1977; Kurachi et al., 1980; Ohkubo et al., 1982). During the activation of factor XI, an internal peptide bond is cleaved in each of the two chains, resulting in factor XI_a, a serine protease composed of two heavy and two light chains, and these four chains are held together by disulfide bonds (Bouma et al., 1977; Kurachi & Davie, 1977). Trypsin also activates factor XI by cleaving the same internal peptide bonds, but this reaction readily occurs in the absence of a surface and HMW kininogen (Mannhalter et al., 1980). When factor XI_a is formed, it in turn activates factor IX in the presence of calcium ions (Fujikawa et al., 1974; DiScipio et al., 1978; Osterud et al., 1978). This latter reaction results in the cleavage of two internal peptide bonds in factor IX and triggers the middle phase of the intrinsic pathway of blood coagulation.

The heavy and light chains of factor XI_a have been separated after mild reduction and S-alkylation. The heavy chains originate from the amino-terminal portion of the molecule and are responsible for the binding of the protein to HMW kininogen. The light chains contain the catalytic activity of the enzyme (van der Graaf et al., 1983). The presence of two catalytic sites per mole of factor XI_a has been demonstrated by titration with antithrombin III (Kurachi & Davie, 1977). More recently, the interaction of the heavy chain of human factor XI_a with factor IX has been reported (Sinha et al., 1985).

Considerable sequence homology has been shown in the amino-terminal portion of factor XI and plasma prekallikrein (Kurachi et al., 1980; Heimark & Davie, 1981). The amino-terminal sequence of the light chain and the region near the catalytic serine residue of factor XI_a have also been published (Koide et al., 1977b; Kurachi et al., 1980). In this paper, the nucleotide sequence of a cDNA coding for human factor XI is reported along with the predicted amino acid sequence. The high degree of amino acid sequence homology of factor XI and plasma prekallikrein is also reported.

Experimental Procedures

Purification of Human Factor XI. The procedure of Heimark and Davie (1981) for the isolation of plasma prekallikrein was employed up to the step of the heparin–agarose column. The fractions from the heparin–agarose column were assayed for factor XI activity (Kurachi & Davie, 1981). The fractions containing factor XI activity were pooled and dialyzed against 20 L of 20 mM Tris-HCl buffer, pH 7.2, containing 70 mM NaCl. The dialyzed sample was stirred for 15 min with 200 mL of CM-Sephadex C-50 that had been equilibrated with the same buffer. The gel was poured into a plastic column (5 × 10 cm) and washed with the buffer until the absorbance at 280 nm reached approximately 0.05. Adsorbed proteins were eluted with a linear gradient composed of 500 mL of 0.07 M NaCl in 20 mM Tris-HCl buffer, pH 7.2, and 500 mL of 0.6 M NaCl in the same buffer. Virtually all of the factor XI clotting activity was eluted in a single broad protein peak in approximately one-third of the gradient. The protein fractions were pooled and dialyzed against 50 mM Tris-HCl buffer, pH 7.5, containing 0.15 M NaCl (TBS). The protein sample was then applied to a HMW kininogen–Sepharose column (20 mL, 2 × 6.5 cm) that had been previously equilibrated with TBS. After the column was washed with TBS, the adsorbed proteins were eluted with 50 mM Tris-HCl, pH 7.5, containing 1 M NaCl. A single, broad protein peak with factor XI activity was obtained. The protein fractions were pooled and concentrated by an Amicon ultrafiltrator with a PM-10 membrane. The final preparation showed one major protein band on SDS–polyacrylamide gel electrophoresis along with a faint contaminant corresponding to IgG. Approximately 3 mg of factor XI was obtained from 4.5 L of fresh frozen human plasma. Molecular weights of 130 000 before reduction and 80 000 after reduction were obtained for the purified factor XI by SDS–polyacrylamide gel electrophoresis (Weber & Osborn, 1969).

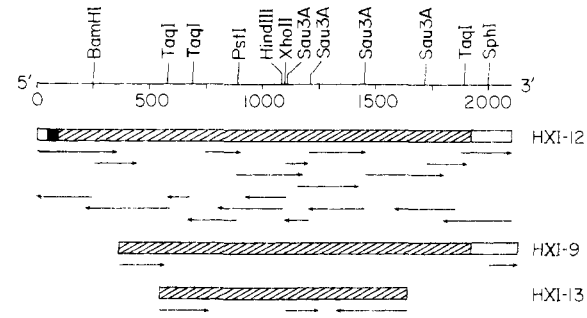


FIGURE 1: Summary of sequencing strategy of the cDNA inserts in λHXI-12, λHXI-13, and λHXI-9. Restriction sites used for the sequence determination are shown. The arrows under the bars indicate the direction and extent of sequence analysis of the respective clones. The open, solid, and slashed bars represent the noncoding region, the leader peptide, and the mature protein, respectively.

Preparation of Antibody to Human Factor XI. Antibody to human factor XI was raised in rabbits by repeated injections of 0.5 mg of the purified protein. A total of 200 mL of blood was withdrawn from two rabbits 5 weeks after the first injection, and the IgG fraction was prepared according to Harboe and Ingild (1973). The rabbit antibody was dialyzed against TBS and passed through a column containing 2 mL of human IgG–Sepharose in order to remove contaminating antibodies to human IgG. The unadsorbed fraction from this column was applied to a 4-mL column containing human factor XI–Sepharose equilibrated with TBS. After the column was washed extensively with 1 L of TBS, the adsorbed antibody was eluted with 0.2 M glycine hydrochloride, pH 2.5. The eluate was mixed immediately with an equal volume of 1 M Tris-HCl buffer, pH 8.0, and the fractions that contained protein were pooled, concentrated, and dialyzed against TBS by using a MicroProDiCon concentrator.

Approximately 0.13 mg of affinity-purified antibody to human factor XI was then radiolabeled with 0.2 mCi of $Na^{125}I$ (Amersham) using Iodogen (Pierce). The antibody had a specific activity of 2×10^9 cpm/mg.

Human IgG was prepared by the method of Harboe and Ingild (1973). Human factor XI–Sepharose and human IgG–Sepharose were prepared by coupling 4 mg of each purified protein to 0.5 g of activated CH-Sepharose (Pharmacia) mixed with 2 mL of Sepharose 4B. HMW kininogen–Sepharose was prepared by coupling 200 mg of the purified protein to 6 g of activated CH-Sepharose according to the manufacturer's instructions.

Screening of λgt11 Library. A human cDNA library cloned into λgt11 phage as described by Kwok et al. (1985) was kindly provided by Dr. Savio Woo. Approximately 4.5 million phage were screened by using affinity-purified antibody by the method of Young and Davis (1983a,b) as modified by Foster and Davie (1984). Antibody containing 1.6×10^6 cpm was used for each 132-mm nitrocellulose filter during the first screen and 5.0×10^5 cpm for each 72-mm filter for subsequent screenings. Positive clones were amplified by the plate lysis method, and the phage were purified by CsCl banding.

Restriction fragments from the isolated inserts were cloned into either M13mp18 or M13mp19 vectors for sequencing by the dideoxy chain termination method (Sanger et al., 1977),

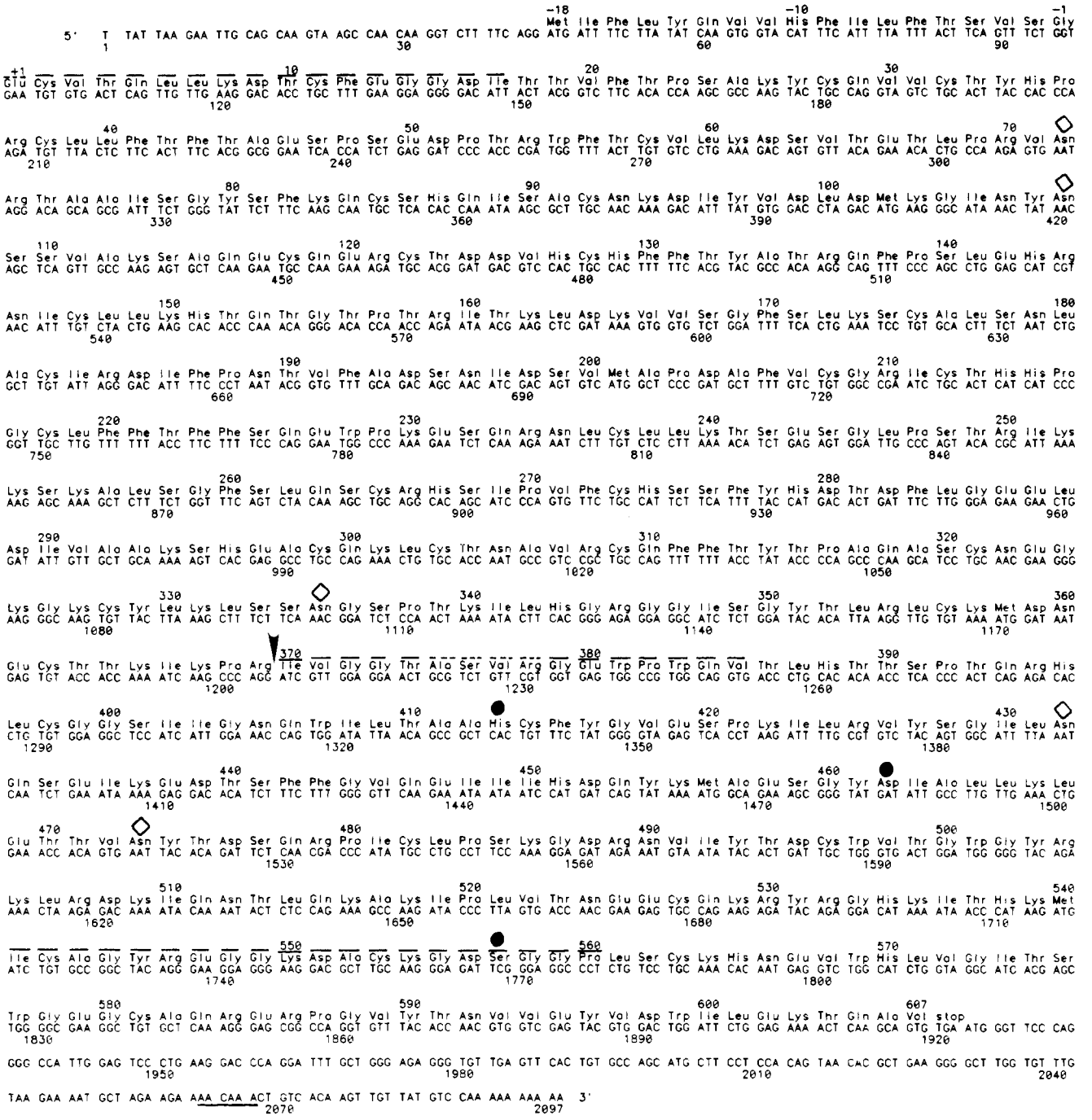


FIGURE 2: Complete nucleotide sequence of the cDNA in λHXI-12 and its predicted amino acid sequence. The amino acid sequence corresponding to the mature protein is numbered from 1 to 607, and the amino acid sequence that corresponds to the leader peptide is represented by the minus numbers in the opposite direction. The residues marked by the overline have been determined by amino acid sequencing (Kurachi & Davie, 1977). The residues with a solid overline agreed and those with a dotted overline disagreed with those predicted by the nucleotide sequence. The Arg–Ile bond shown by the arrowhead is the site of cleavage by factor XII_a (or factor XII). The Asn residues shown by open diamonds are the potential glycosylation sites. The catalytic triad of His, Asp, and Ser is identified by solid circles. The nucleotide sequence that is underlined is the potential recognition signal for polyadenylation or processing.

incorporating [^{35}S]dATP for sequencing on a 6% polyacrylamide buffer–gradient gel system (Yoshitake et al., 1985). All endonucleases were purchased from BRL except for *Xho*II which was a product of New England Biolabs.

Results and Discussion

A λgt11 expression library containing cDNAs prepared from human liver mRNA was screened for factor XI by employing an affinity-purified antibody. Five positive clones were isolated and plaque purified. Phage DNA was isolated from the positive clones, and the cDNA insert was released after digestion with endonuclease *Eco*RI. These phage were found to contain cDNA inserts of the following size: λHXI-12, 2.1 kb; λHXI-13, 1.1 kb; λHXI-9, 1.8 kb; λHXI-10, 1.1 kb; and λHXI-11, 1.1 kb. The cDNA inserts in λHXI-12, λHXI-13, and λHXI-9 were then subjected to nucleotide sequence analysis by the dideoxy chain termination method of Sanger et al. (1977) to establish their terminal sequences. The insert in λHXI-12 was found to be a near full-length cDNA coding for human factor XI. It was then subcloned into the *Eco*RI site of plasmid pBR322 in order to establish its entire nucleotide sequence. Sequencing was performed on both strands of the cDNA insert in λHXI-12, except for nucleotides 440–743 and 1451–1568 where the analysis was done in only

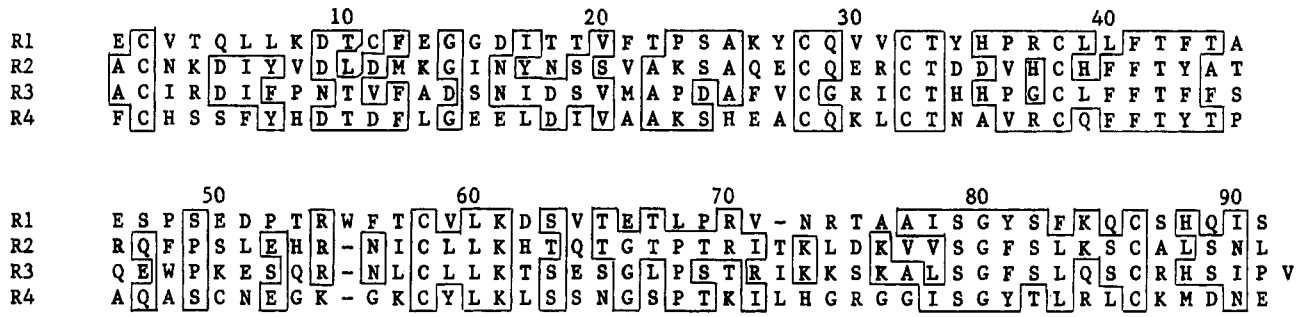


FIGURE 3: Alignment for the four tandem repeat sequences in the heavy chain of factor XI_a. Gaps were inserted in two places for maximal alignment. The positions that have two or more identical residues are boxed. The numbers indicate the length of the peptide for each repeat. R1, R2, R3, and R4 represent the first, second, third, and fourth repeats, respectively.

one direction. Sequencing of the complementary strand for these regions was performed with the cDNA insert in λHXI-13 or λHXI-9. A summary of the sequencing strategy and restriction mapping is shown in Figure 1.

The complete nucleotide sequence of the insert in λHXI-12 is presented in Figure 2. This cDNA contained 2097 nucleotides including 43 nucleotides of 5′ noncoding sequence, 54 nucleotides coding for a leader peptide of 18 amino acids, 1821 nucleotides coding for the mature protein, a stop codon of TGA, 166 nucleotides of 3′ noncoding sequence, and a poly(A) tail of 10 nucleotides. A poly(A) tail of 53 bases was observed in the cDNA insert in λHXI-9. The typical polyadenylation or processing sequence of AATAAA (Proudfoot & Brownlee, 1981) was not present in λHXI-12 or λHXI-9 near the 3′ noncoding region. A sequence of AACAAA that is similar to AATAAA was present 21 nucleotides upstream from the poly(A) tail. Accordingly, it seems likely that this sequence is the recognition site for polyadenylation or processing. Sequences of AAGAAA that are also similar to AATAAA are present 27 and 40 nucleotides upstream from the poly(A) tail. This sequence, however, has been shown to inhibit the cleavage step preceding polyadenylation (Montell et al., 1983) and thus is probably not involved in the processing of the mRNA for factor XI.

The predicted amino acid sequence for factor XI reveals the presence of a leader peptide which is rich in hydrophobic amino acids (12 of 18 residues). This leader peptide is typical of secreted proteins and is removed by the cleavage of a Gly–Glu peptide bond by signal peptidase during biosynthesis and secretion across the rough endoplasmic reticulum (Blobel et al., 1979). A stop codon (TAA) was present at nucleotides 5–7, indicating that the methionine codon at nucleotides 44–46 was the correct initiation site. The mature protein that circulates in plasma contains 607 amino acid residues in each chain. Thus, the dimer of factor XI is composed of 1214 residues with the following composition: Asp_{56}, Asn_{42}, Thr_{104}, Ser_{98}, Glu_{62}, Gln_{52}, Gly_{84}, Pro_{48}, Ala_{60}, Val_{70}, $^1/_2$-Cys_{72}, Met_{10}, Leu_{88}, Ile_{72}, Phe_{52}, Tyr_{42}, Trp_{20}, Lys_{82}, His_{42}, and Arg_{58}. A molecular weight of 135 979 was calculated from the composition for the two polypeptide chains. Addition of 5% carbohydrate gives an approximate molecular weight of 143 000 for human factor XI. This is in fair agreement with the reported values of 130 000 (Kurachi & Davie, 1977) and 160 000 (Bouma & Griffin, 1977) as determined by SDS–polyacrylamide gel electrophoresis. The existence of a large excess of basic amino acids (182 residues) compared to the acidic residues (118) in the monomer of factor XI is reflected by the basic nature of the protein. This is further evident from the fact that factor XI does not adsorb to DEAE-Sephadex at pH 7.5 with a low ionic strength buffer (5 $m\Omega^{-1}$ conductivity) (Fujikawa & Davie, 1981).

During the activation of factor XI, a single internal peptide bond in each polypeptide chain is cleaved by factor XII_a (or factor XII). This converts factor XI into factor XI_a which is composed of two heavy chains and two light chains, and all four chains are held together by disulfide bonds. The amino-terminal sequence of the light chain of human factor XI_a has been reported as Ile-Val-Gly-Gly- (Kurachi & Davie, 1977). Therefore, the cleavage site in factor XI by factor XII_a occurs at the peptide bonds between each Arg-369 and Ile-370. Each heavy chain in factor XI_a consists of 369 amino acid residues starting with the amino-terminal sequence of Glu-Cys-Val-Thr- and ending with the carboxyl-terminal sequence of Ile-Lys-Pro-Arg. The presence of Pro-368 adjacent to Arg-369 could explain the previous failure in identifying the carboxyl-terminal residue of the heavy chain of factor XI_a by digestion with carboxypeptidase B (K. Kurachi, unpublished results). The light chain contains 238 amino acids, starting with Ile-Val-Gly-Gly- and ending with Thr-Gln-Ala-Val.

Four repeated sequences are present in the heavy chain of factor XIa when an alignment was made by the $^1/_2$-Cys residues (Figure 3). Each tandem repeat consists of 90 (or 91) amino acids and includes 6 conserved $^1/_2$-Cys residues. Although the pairings for the disulfide bonds in factor XI have not been determined, it is likely that the six $^1/_2$-Cys residues form three intrachain disulfide bonds within each of the four repeats. Consequently, four separate domains are probably present in each of the two identical heavy chains of factor XI. The identity in these repeats ranges from 23% to 34%.

Out of 36 $^1/_2$-Cys residues, 34 in the monomer of factor XI match perfectly with those in plasma prekallikrein (Figures 4 and 5). In the heavy chain of factor XI_a, the first and fourth repeats have seven $^1/_2$-Cys residues, while six $^1/_2$-Cys residues are present in the second and third repeats. Six of the seven $^1/_2$-Cys residues in the first and fourth repeats are located in the same positions as those in the second and third repeats, suggesting all four domains are similar. In the fourth repeat of plasma prekallikrein, one extra intrachain disulfide bond is present between $^1/_2$-Cys-321 and $^1/_2$-Cys-326 (employing the factor XI numbering). This extra disulfide bond is not present in factor XI, since $^1/_2$-Cys-326 in plasma prekallikrein has been replaced by Gly. The two polypeptide chains in human factor XI are known to be connected by one or more disulfide bonds (Bouma & Griffin, 1977; Kurachi & Davie, 1977). Therefore, the two unmatched $^1/_2$-Cys residues at position 11 in the first repeat and position 321 in the fourth repeat are very likely involved in the interchain disulfide bridge between the two monomers. This assumes the remaining $^1/_2$-Cys residue 362 in the connecting region forms a disulfide bond with $^1/_2$-Cys-113 in the light chain or catalytic domain (Figures 4 and 5). In the catalytic domain of the molecule, it is highly likely that eight $^1/_2$-Cys residues form disulfide

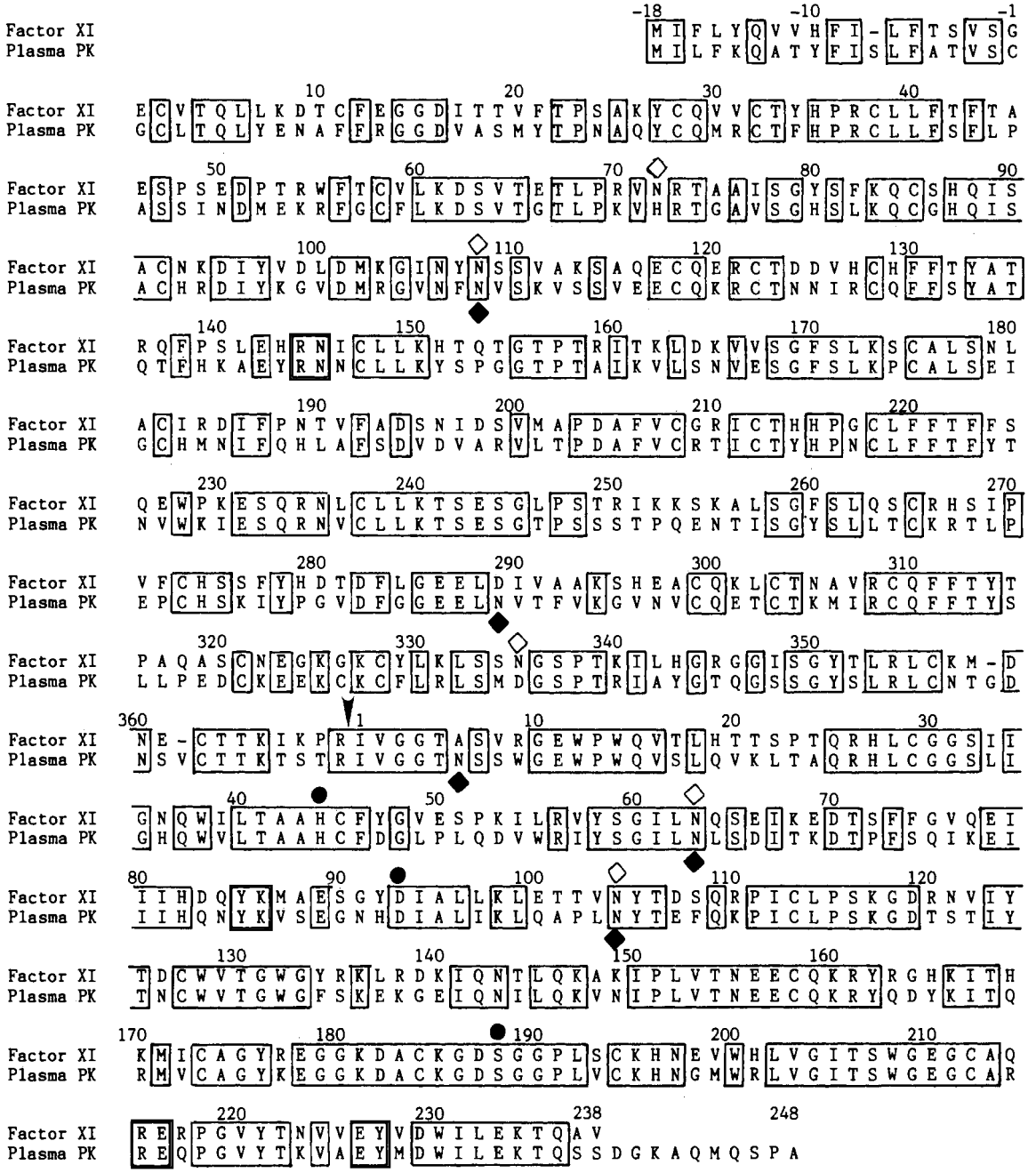
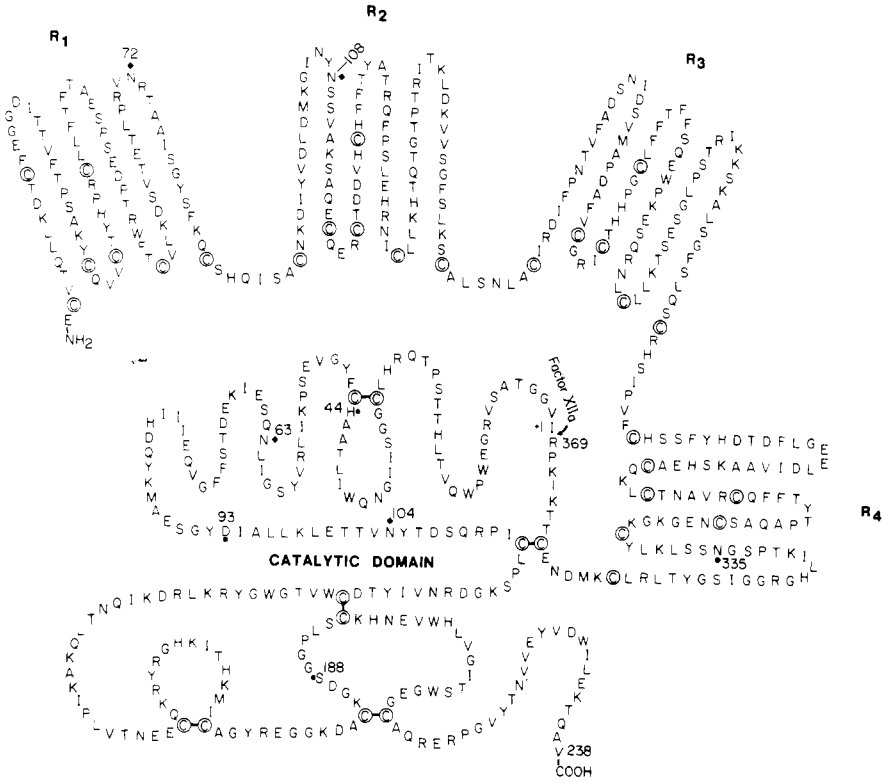


FIGURE 4: Comparison of the amino acid sequence of human factor XI and plasma prekallikrein. The alignment was made by inserting two gaps in the factor XI sequence following residues 358 and 361. The amino acid sequence corresponding to the mature proteins is numbered from 1 to 369 for the heavy chain. The numbering system differs from that in Figure 2 in that the catalytic or light chain of factor XI_a and of plasma kallikrein are numbered 1–238 and 1–248, respectively. The sequence that corresponds to the leader peptide is represented by minus numbers in the opposite direction. The positions that have identical residues in both proteins are boxed. The Asn residues marked by the solid diamonds in prekallikrein are the glycosylation sites that have been proven by amino acid sequencing. The Asn residues marked by the open diamonds in factor XI are potential glycosylation sites. The Arg–Ile bond marked by the arrowhead is the cleavage site catalyzed by factor XII_a (or factor XII). The catalytic triad of His, Asp, and Ser is identified by the solid circles. The sequence of plasma prekallikrein was taken from Chung et al. (1986).

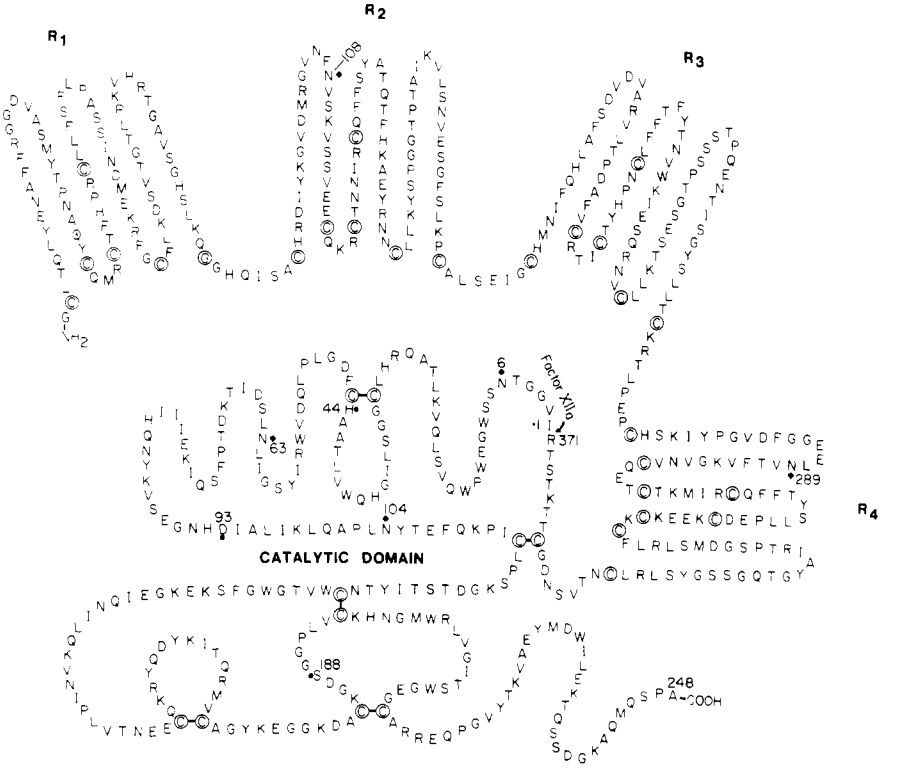
bonds in a manner analogous to other serine proteases, such as plasmin (Wieman 1977), trypsin, and chymotrypsin (Figure 5). In these proteins, the disulfide bonds have been established by chemical methods. If two interchain disulfide bonds occur between the monomers in factor XI, two different structures are possible for the dimer. These are either a *parallel* or an *antiparallel* structure.

The catalytic triad, including His-44, Asp-93, and Ser-188, is present in the light chain of factor XIa (Figures 4 and 5). Asp-182 which is probably located at the bottom of the substrate binding pocket indicates that factor XI_a is a typical trypsin-like serine protease. These structural features are characteristic of the trypsin family of proteases and are consistent with the substrate specificity of factor XI_a. During the coagulation process, factor XI_a catalyzes the activation of factor IX by the cleavage of an internal Arg–Ala bond and an Arg–Ile bond (Fujikawa et al., 1974; Lindquist et al., 1978; DiScipio et al., 1978; Katayama et al., 1979; Kurachi & Davie, 1982). Factor XI_a also hydrolyzes arginyl ester or arginyl-amide bonds, such as those in tosylarginyl methyl ester (Kurachi et al., 1980) and pyro-Glu-Pro-Arg-pNA (Scott et al., 1984).

Short amino acid sequences for factor XIa have been determined in three different regions of the molecule (Kurachi



Factor XI



Plasma Prekallikrein

FIGURE 5: Amino acid sequence and tentative structures of the monomer of factor XI (top) and plasma prekallikrein (bottom). The disulfide bonds in the light chains are placed by analogy with the B chain of plasmin (Wieman, 1977). The cleavage sites by factor XII_a (or factor XII) are shown by curved arrows. The glycosylation sites are marked by the solid diamonds, and the catalytic triad of His, Asp, and Ser is identified by the solid circles. The numbering system is the same as that in Figure 4.

& Davie, 1977). These sequences correspond to position 1 through Ile-17, to Ile-370 through Val-385, and to Ile-541 through Pro-560 (Figure 2). These data are consistent with the present results except for the following amino acids: Gly was reported at position 1 by protein sequencing instead of Glu [Glu was also identified in bovine factor XI (Kurachi et al., 1980)], and Gln was originally reported at position 11 (Kurachi & Davie, 1977) but corrected later to $^1/_2$-Cys

(Kurachi et al., 1980). Also the sequence of Val-Ala-?-His at positions 375–378 was reported instead of Ala-Ser-Val-Arg. The latter amino acid sequence was also predicted by the nucleotide sequence in λHXI-13. The reasons for the last discrepancy are not known.

Five potential Asn-linked carbohydrate attachment sites are present in the monomer of factor XI. These are located at residues 72, 108, and 335 in the heavy chain and 63 and 104 in the light chain (Figures 4 and 5). Each is present in the consensus sequence of Asn-X-Ser or Asn-X-Thr. Three other Asn residues (145, 235, and 360 in the heavy chain) are followed by X-Cys and are also potential glycosylation sites that have been identified in protein C (Stenflo et al., 1975) and von Willebrand factor (Titani et al., 1985). Plasma prekallikrein has five Asn-linked carbohydrate attachment sites including two in the heavy chain and three in the light chain (Figures 4 and 5). The glycosylation of these five Asn residues has been demonstrated by protein sequencing (Chung et al., 1986). Three of the five potential Asn-linked carbohydrate attachment sites in factor XIa are located at positions identical with those in plasma kallikrein, including Asn-108 in the heavy chain and Asn-63 and -104 in the light chain. The carbohydrate content of factor XI, however, is 5% and is much lower than the highly homologous plasma prekallikrein (15.5%). Accordingly, glycosylation is probably absent in some of the potential carbohydrate attachment sites, or the lengths of the carbohydrate chains are shorter in human factor XI.

The amino acid sequences of human factor XI and plasma prekallikrein are highly homologous (58% identity) (Figures 4 and 5). The leader peptide of factor XI is shorter by one amino acid than that of plasma prekallikrein. These leader peptides show 50% amino acid identity. The polypeptide chain of the factor XI monomer (607 amino acids) is also slightly shorter (12 residues) than plasma prekallikrein. This is due to the presence of 10 additional amino acids at the carboxyl-terminal end of plasma prekallikrein and 2 extra amino acids in the connecting region located between the fourth repeat and the light chain (Figure 4). Overall, there is a very high degree of sequence identity between both proteins throughout the whole molecule. The degree of homology, which includes similar amino acids such as Ile/Val, Leu/Val, Leu/Ile, Thr/Ser, and Tyr/Phe, is 67% between the two proteins. An extremely high degree of identity (81%) is present, starting with Ile-150 in the light chain and extending to the carboxyl-terminal end of the catalytic domain (Figure 4). This region includes the substrate binding pocket and the active-site serine residue.

In spite of the high homology between factor XI and plasma prekallikrein, the biological functions of these two proteins are distinct. The major function of factor XI_a is in the activation of factor IX during blood coagulation, whereas kallikrein is incapable of activating factor IX. Plasma kallikrein also has a somewhat broader substrate specificity. For instance, plasma kallikrein is the principal enzyme that activates factor XII in the presence of an anionic surface and HMW kininogen. In contrast, factor XI_a has weak activity in the activation of factor XII (Davie et al., 1979; Griffin & Cochrane, 1979; Colman, 1984). Plasma kallikrein releases the vasoactive peptide bradykinin from HMW kininogen (Kato et al., 1981) and is also involved in the intrinsic pathway of fibrinolysis where it is capable of activating prourokinase (Ichinose et al., 1985). Recently, factor XI_a also has been shown to release bradykinin from HMW kininogen (Scott et al., 1985). The efficiency of factor XI_a in the release of bradykinin, however, is substantially less than that of kallikrein.

Factor XI and plasma prekallikrein each circulate as a complex with HMW kininogen in blood. The competitive binding of these two proteins to the light chain of HMW kininogen has been reported (Thompson et al., 1979). This suggests that these two proteins share the same binding site in the HMW kininogen molecule. The binding constant of prekallikrein, however, is higher than that of factor XI (Thompson et al., 1979; Bouma et al., 1983). These results are consistent with the present experiments in which an affinity column of HMW kininogen has been employed to separate the two proteins. In these experiments, a higher salt concentration was required for the elution of prekallikrein than for factor XI.

The isolation of the cDNAs for factor XI and plasma prekallikrein (Chung et al., 1986) makes it possible to study and compare the organization of the genes for these two closely related proteins. It will be of particular interest to identify the location of the introns in the genes for the two proteins and determine whether or not they occur between the four homologous tandem repeats that are present in each of these two proteins.

ACKNOWLEDGMENTS

We thank Drs. Savio Woo and Vincent Kidd, Baylor University, for a gift of human cDNA library and Drs. Steven Leytus and Donald Foster for helpful discussions. We also thank Jeffery Harris for technical assistance.

REFERENCES

Blobel, G., Walter, P., Chang, C. N., Goldman, B. M., Erickson, A. H., & Lingappa, R. (1979) *Soc. Exp. Biol. Symp. 33*, 9–36.

Bouma, B. N., & Griffin, J. H. (1977) *J. Biol. Chem. 252*, 6432–6437.

Bouma, B. N., Vlooswijk, R. A. A., & Griffin, J. H. (1983) *Blood 62*, 1123–1131.

Colman, R. W. (1984) *J. Clin. Invest. 73*, 1249–1253.

Chung, D. W., Fujikawa, K., McMullen, B. A., & Davie, E. W. (1986) *Biochemistry* (preceding paper in this issue).

Davie, E. W., Fujikawa, K., Kurachi, K., & Kisiel, W. (1979) *Adv. Enzymol. Relat. Areas Mol. Biol. 48*, 277–318.

DiScipio, R. G., Kurachi, K., & Davie, E. W. (1978) *J. Clin. Invest. 61*, 1528–1538.

Foster, D. C., & Davie, E. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4766–4770.

Fujikawa, K., & Davie, E. W. (1981) *Methods Enzymol. 80*, 198–211.

Fujikawa, K., Legaz, M. E., Kato, H., & Davie, E. W. (1974) *Biochemistry 13*, 4508–4516.

Griffin, J. H., & Cochrane, C. G. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2554–2558.

Griffin, J. H., & Cochrane, C. G., (1979) *Semin. Thromb. Hemostasis 5*, 254–273.

Harboe, N., & Ingild, A. (1973) *Scand. J. Immunol. 2, Suppl. 1*, 161–164.

Heinmark, R. L., & Davie, E. W. (1981) *Methods Enzymol. 80*, 157–172.

Ichinose, A., Fujikawa, K., & Suyama, T. (1986) *J. Biol. Chem. 261*, 3486–3489.

Katayama, K., Ericsson, L. H., Enfield, D. L., Walsh, K. A., Neurath, H., Davie, E. W., & Titani, K. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4990–4994.

Kato, H., Nagasawa, S., & Iwanaga, S. (1981) *Methods Enzymol. 80*, 172–198.

Koide, T., Kato, H., & Davie, E. W. (1976) *Methods Enzymol. 45*, 65–73.

Koide, T., Kato, H., & Davie, E. W. (1977a) *Biochemistry 16*, 2279–2286.

Koide, T., Hermodson, M. A., & Davie, E. W. (1977b) *Nature (London) 266*, 729–730.

Kurachi, K., & Davie, E. W. (1977) *Biochemistry 16*, 5831–5839.

Kurachi, K., & Davie, E. W. (1981) *Methods Enzymol. 80*, 211–220.

Kurachi, K., & Davie, E. W. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6461–6464.

Kurachi, K., Fujikawa, K., & Davie, E. W. (1980) *Biochemistry 19*, 1330–1338.

Kwok, S. C., Ledley, F. D., DiLella, A. G., Robson, K. J., & Woo, S. L. C. (1985) *Biochemistry 24*, 556–561.

Lindquist, P. A., Fujikawa, K., & Davie, E. W. (1978) *J. Biol. Chem. 253*, 1902–1909.

Mannhalter, C., Schiffman, S., & Jocobs, A. (1980) *J. Biol. Chem. 255*, 2667–2679.

Montell, G., Fisher, E. F., Caruthers, M. H., & Berk, A. J. (1983) *Nature (London) 305*, 600–605.

Ohkubo, I., Fujikawa, K., & Kurachi, K. (1982) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 41*, 656.

Osterud, B., Bouma, B. N., & Griffin, J. H. (1978) *J. Biol. Chem. 253*, 5946–5951.

Proudfoot, N., & Brownlee, G. G. (1981) *Nature (London) 252*, 359–362.

Rapaport, S. I., Proctor, R. R., Patch, M. J., & Yettra, M. (1961) *Blood 18*, 149–165.

Ratnoff, O. D., & Saito, H. (1979) *Curr. Top. Hematol. 2*, 1–57.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Scott, C. F., Sinha, D., Seaman, F. S., Walsh, P. N., & Colman, R. W. (1984) *Blood 63*, 42–50.

Scott, C. F., Silver, L. D., Purdon, A. D., & Colman, R. W. (1985) *J. Biol. Chem. 260*, 10856–10863.

Sinha, D., Koshy, A., Seaman, F. S., & Walsh, P. N. (1985) *J. Biol. Chem. 260*, 10714–10719.

Stenflo, J., & Fernlund, P. (1982) *J. Biol. Chem. 259*, 12180–12190.

Thompson, R. E., Mandle, R., Jr., & Kaplan, A. P. (1977) *J. Clin. Invest. 60*, 1376–1380.

Thompson, R. E., Mandle, R., Jr., & Kaplan, A. P. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4862.

Titani, K., Kumar, S., Takio, K., Ericsson, L. H., Wade, R. D., Ashida, K., Walsh, K. A., Chopek, M. W., Sadler, J. E., & Fujikawa, K. (1986) *Biochemistry* (in press).

van der Graaf, F., Greengard, J. S., Bouma, B. N., Kerbiriou, D. M., & Griffin, J. H. (1983) *J. Biol. Chem. 258*, 9669–9675.

Weber, K., & Osborn, M. (1969) *J. Biol. Chem. 244*, 4406–4412.

Wieman, B. (1977) *Eur. J. Biochem. 76*, 129–137.

Yoshitake, S., Schach, B. G., Foster, D. C., Davie, E. W., & Kurachi, K. (1985) *Biochemistry 24*, 3736–3750.

Young, R. A., & Davis, R. W. (1983a) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1194–1198.

Young, R. A., & Davis, R. W. (1983b) *Science (Washington, D.C.) 222*, 778–782.