

# Human Tissue Factor: cDNA Sequence and Chromosome Localization of the Gene<sup>†</sup>

Eleonora M. Scarpati,<sup>‡,§,||</sup> Duanzhi Wen,<sup>‡,§,||</sup> George J. Broze, Jr.,<sup>§</sup> Joseph P. Miletich,<sup>§</sup> Robert R. Flandermeyer,<sup>‡</sup> Ned R. Siegel,<sup>#</sup> and J. Evan Sadler<sup>\*‡,§,||</sup>

Howard Hughes Medical Institute and Departments of Medicine and Biological Chemistry, Washington University School of Medicine, St. Louis, Missouri 63110, Department of Obstetrics, Gynecology, and Reproductive Science, University of California, San Francisco, California 94143, and Monsanto Company, St. Louis, Missouri 63198

Received June 3, 1987; Revised Manuscript Received June 25, 1987

**ABSTRACT:** A human placenta cDNA library in  $\lambda$ gt11 was screened for the expression of tissue factor antigens with rabbit polyclonal anti-human tissue factor immunoglobulin G. Among 4 million recombinant clones screened, one positive,  $\lambda$ HTF8, expressed a protein that shared epitopes with authentic human brain tissue factor. The 1.1-kilobase cDNA insert of  $\lambda$ HTF8 encoded a peptide that contained the amino-terminal protein sequence of human brain tissue factor. Northern blotting identified a major mRNA species of 2.2 kilobases and a minor species of  $\sim$ 3.2 kilobases in poly(A)<sup>+</sup> RNA of placenta. Only 2.2-kilobase mRNA was detected in human brain and in the human monocytic U937 cell line. In U937 cells, the quantity of tissue factor mRNA was increased severalfold by exposure of the cells to phorbol 12-myristate 13-acetate. Additional cDNA clones were selected by hybridization with the cDNA insert of  $\lambda$ HTF8. These overlapping isolates span 2177 base pairs of the tissue factor cDNA sequence that includes a 5'-noncoding region of 75 base pairs, an open reading frame of 885 base pairs, a stop codon, a 3'-noncoding region of 1141 base pairs, and a poly(A) tail. The open reading frame encodes a 33-kilodalton protein of 295 amino acids. The predicted sequence includes a signal peptide of 32 or 34 amino acids, a probable extracellular factor VII binding domain of 217 or 219 amino acids, a transmembrane segment of 23 amino acids, and a cytoplasmic tail of 21 amino acids. There are three potential glycosylation sites with the sequence Asn-X-Thr/Ser. The 3'-noncoding region contains an inverted Alu family repetitive sequence. The tissue factor gene was localized to chromosome 1 by hybridization of the cDNA insert of  $\lambda$ HTF8 to flow-sorted human chromosomes.

**T**issue factor is an integral plasma membrane glycoprotein that binds to blood coagulation factor VII or factor VIIa. This interaction increases the rate of activation of factor X by factor VII/VIIa, initiating blood clotting by the extrinsic pathway (Straub & Duckert, 1961; Silverberg et al., 1977). In addition, the tissue factor-factor VIIa complex can activate factor IX, so that tissue factor may also promote clotting by recruiting the intrinsic pathway (Østerud & Rapaport, 1977). Tissue factor is found in many cell types that are not exposed directly to flowing blood and is particularly abundant in the parenchyma of brain, lung, and placenta (Williams, 1964, 1966). Thus, injuries that disrupt the endothelium will bring factor VII into contact with tissue factor to cause blood clotting and restore hemostasis.

The vascular endothelium does not normally express tissue factor activity but can be induced to do so by stimulation with

thrombin (Brox et al., 1984), phorbol esters (Lyberg et al., 1983; Stern et al., 1984), phytohemagglutinin, endotoxin (Lyberg et al., 1983), interleukin 1 (Nawroth et al., 1986), or tumor necrosis factor (Nawroth & Stern, 1986). Consequently, endothelial cells may be able to regulate clotting reactions at their surface by modulating the expression of tissue factor activity. Blood monocytes also express tissue factor in response to many of these stimuli (Rivers et al., 1975; Østerud and Bjørklid, 1982; Lyberg et al., 1982). These phenomena may contribute to the pathophysiology of disseminated intravascular coagulation and venous thromboembolism in some disease states.

Little is known about the structure of tissue factor. It has been purified to apparent homogeneity from bovine (Bach et al., 1981) and human brain (Broze et al., 1985; Tanaka et al., 1985; Guha et al., 1986). Human tissue factor appears to be a glycoprotein containing a single polypeptide chain with an apparent  $M_r$  of 44 000 upon sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Bach et al., 1981; Broze et al., 1985; Tanaka et al., 1985; Guha et al., 1986).

We report here the cloning and essentially full-length nucleotide sequence of cDNA isolates for human tissue factor. A preliminary report of this work has been presented (Scarpatti et al., 1987). The isolation of partial cDNA clones for human tissue factor has been reported by Morrissey et al. (1987a).

## EXPERIMENTAL PROCEDURES

**Materials.** Goat anti-rabbit immunoglobulin G (IgG) conjugated to horseradish peroxidase was from Bio-Rad. RNA size standards were from Bethesda Research Laboratories. 7-Deoxy-7-deazaguanosine 5'-triphosphate was from Boehringer Mannheim. <sup>32</sup>P-Labeled deoxyribonucleotides and

\* This work was supported in part by Research Grants HLBI 14147 (Specialized Center for Research in Thrombosis) and HLBI 34462 (to G.J.B.) from the National Institutes of Health, Bethesda, MD, and by a Monsanto-Washington University Biomedical Research Grant. E.M.S. was supported in part by the Consejo Nacional de Investigaciones Científicas y Tecnológicas, Caracas, Venezuela, and by the Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO.

<sup>†</sup> Correspondence should be addressed to this author at Howard Hughes Medical Institute.

<sup>‡</sup> Howard Hughes Medical Institute, Washington University School of Medicine.

<sup>§</sup> Department of Medicine, Washington University School of Medicine.

<sup>||</sup> Department of Biological Chemistry, Washington University School of Medicine.

<sup>#</sup> Department of Obstetrics, Gynecology, and Reproductive Science, University of California.

<sup>\*</sup> Monsanto Co.

deoxyadenosine 5'-([ $\alpha$ -<sup>35</sup>S]thiotriphosphate) were from Amersham Radiochemicals.

**Antibody to Human Tissue Factor.** Thirty micrograms of purified human brain tissue factor (Broze et al., 1985) in complete Freund's adjuvant was injected into the popliteal lymph node of each hind leg of a 2-kg male New Zealand white rabbit (Sigel et al., 1983). After 3 weeks, additional purified tissue factor (60  $\mu$ g total) in incomplete Freund's adjuvant was injected into the thighs and into multiple intradermal sites on the back (Vaitukaitus, 1981). Every 2-3 weeks thereafter, blood was drawn from the rabbit and polyclonal anti-tissue factor IgG isolated with ammonium sulfate precipitation and chromatography on DEAE-Affi-Gel Blue as described by the manufacturer (Bio-Rad, Richmond, CA). Anti-tissue factor activity was detected by using a specific two-stage bioassay (Broze et al., 1985).

**Isolation of cDNA Clones.** The human placenta cDNA library in  $\lambda$ gt11, methods for screening with antibodies, preparation and use of cDNA restriction fragment probes, and preparation of  $\lambda$ -phage DNA were as described previously (Ye et al., 1987). The rabbit polyclonal anti-human tissue factor serum was diluted 1/125 for screening of the cDNA library.

**Characterization of cDNA Clones by Epitope Selection.** Recombinant proteins induced by isopropyl  $\beta$ -thiogalactoside from positive  $\lambda$ -phage isolates or from control  $\lambda$ gt11 were employed for the affinity purification of antibody from rabbit polyclonal anti-human tissue factor IgG by the method of Weinberger et al. (1985), with modifications described previously (Ye et al., 1987). The eluted and dialyzed antibodies were tested for recognition of authentic human brain tissue factor in concentrations from 0.1 to 50 ng by using the assay employed for cDNA library screening.

**Chromosome Localization of the Tissue Factor Gene.** The cDNA insert of  $\lambda$ HTF8 was labeled with [ $\alpha$ -<sup>32</sup>P]dCTP by nick translation to  $\geq 10^8$  cpm/ $\mu$ g and hybridized to spot blots of flow-sorted human chromosomes (Lebo et al., 1984). The blots were washed, and the signals were detected by autoradiography as previously described (Lebo et al., 1984).

**Other Methods.** Preparation of RNA, Northern and Southern blotting, subcloning, and DNA sequencing were performed as described previously (Ye et al., 1987; Wen et al., 1987). Nucleotide and protein sequences were analyzed and compared to the National Biomedical Research Foundation Protein Sequence Database (Georgetown University, Washington, DC, release 12.0, March 17, 1987) and the Genbank Genetic Sequence Data Bank (BBN Laboratories Inc., Cambridge, MA, release 48.0, Feb 16, 1987) as described previously (Wen et al., 1987). Human brain tissue factor was purified, and the amino-terminal protein sequence was determined as described previously (Broze et al., 1985).

## RESULTS AND DISCUSSION

**Screening of a Placenta  $\lambda$ gt11 cDNA Library and Characterization of Recombinant Proteins.** Approximately 4 million recombinant clones were screened with rabbit anti-human tissue factor antibody to obtain three positives,  $\lambda$ HTF8 (1.0 kb),  $\lambda$ HTF15 (0.4 kb), and  $\lambda$ HTF16 (2.0 kb). The cDNA inserts did not cross hybridize, suggesting that these isolates encoded unrelated proteins. The immunological relationship between the proteins expressed by  $\lambda$ HTF8,  $\lambda$ HTF16, and authentic human brain tissue factor was tested by the method of "epitope selection" (Weinberger et al., 1985). Proteins induced by isopropyl  $\beta$ -thiogalactoside in *Escherichia coli* infected by these  $\lambda$  phage or by  $\lambda$ GT11 were bound to nitrocellulose filters and used for affinity purification of the rabbit anti-human tissue factor antibody. Isolate  $\lambda$ HTF8 yielded

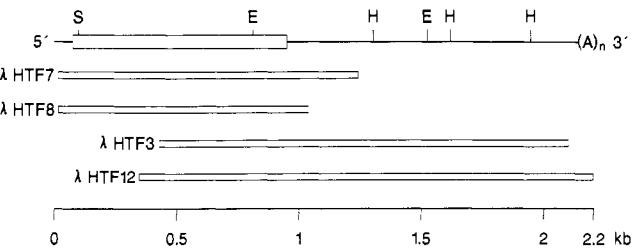


FIGURE 1: Restriction map of tissue factor cDNA isolates. The 5' and 3' ends of the restriction map are labeled. Selected restriction sites are shown: E, EcoRI; H, HindIII; S, SmaI. The thin segments indicate noncoding sequences, and the thick segment indicates the open reading frame that encodes tissue factor. The portion of the sequence contained in each of the cDNA isolates,  $\lambda$ HTF7,  $\lambda$ HTF3, and  $\lambda$ HTF12, is shown by the thin open bars. The scale is in kilobases (kb).

antibody that could recognize as little as 1.5 ng of tissue factor but did not bind to the other cDNA isolates. The antibody eluted from  $\lambda$ HTF16 or from  $\lambda$ gt11 did not detect over 50 ng of tissue factor and also did not bind to the proteins expressed by the other isolates (data not shown).

**Nucleotide Sequence and Predicted Protein Sequence of Tissue Factor cDNA Isolates.** Additional cDNA clones were isolated by hybridization with the insert of  $\lambda$ HTF8. A total of 35 positives were obtained after screening of 1.25 million recombinant clones. Four of these were characterized by restriction mapping and DNA sequencing (Figure 1). The composite nucleotide sequence of these isolates was determined completely on both strands and in most regions on multiple independent isolates. Together they span 2.2 kb. The first ATG codon occurs at nucleotide 76 in the sequence GACATGG, which differs at two minor positions from the proposed consensus sequence for initiation by eukaryotic ribosomes, ACCATGG (Kozak, 1986). This proposed initiation codon is followed by an open reading frame of 885 nucleotides, a stop codon, 1141 nucleotides of 3'-noncoding sequence, and a poly(A) tail of 73 nucleotides. There is a single processing or polyadenylation signal of AATAAA (Proudfoot & Brownlee, 1981) 24 nucleotides before the poly(A) tail. The 3'-noncoding region is rich in A + T residues. This is a feature of the RNA sequences of many other inflammatory mediators that may promote the rapid degradation of these mRNA species (Caput et al., 1986; Shaw & Kamen, 1986). There is a complete inverted Alu family repetitive sequence within the 3'-noncoding region (Deininger et al., 1981) (Figure 2). The Alu repeat is flanked by 11 nucleotide palindromic direct repeats, GAAAAGCTTT, that contain HindIII sites.

Proof that these cDNA inserts encode tissue factor was obtained by comparison of the amino acid sequence deduced from the cDNA sequence with that determined by protein sequencing. The discrepancy between the previously reported amino acid sequence (Broze et al., 1985) and that shown in Figure 2 is explained by heterogeneity at the amino terminus of the protein isolated from brain (Table I). The low recovery of PTH-asparagine for position 11 probably reflects glycosylation. The low yields of serine and threonine are characteristic of the reduced efficiency with which these PTH-amino acids are recovered. Analysis of the repetitive yields obtained for the first 19 cycles indicates that  $\sim 58\%$  (64 pmol) of the protein began with Ser-1 and  $\sim 42\%$  (47 pmol) began with Thr-3, with an average repetitive yield of  $\sim 87\%$ .

The 32 amino acids that precede the amino-terminal Ser-1 of brain tissue factor appear to constitute a cleavable signal peptide. The length of the charged amino-terminal segment is longer than in many other signal peptides, but the length and composition of the uncharged hydrophobic segment are

GGGGAGCCTTCAGCCCAACCTCCCCAGCCCCACGGGCACCGAACCCGCTCGATCTCGCCGCAACTGGTAGACATGGAG  
 M E -31

ACCCCTGCCTGGCCCCGGTCCCGCGCCCCGAGACCGCCGCTCGACGCTCCTGCTGGCTGGTCTCGCCAGGTGGCCGGCGCT  
 T P A W P R V P R P E T A V A R T L L L G W V F A Q V A G A 171 -1

TCAGGCACTACAAATACTGTGGCAGCATATAATTAACTTGAATCAACTAATTCAAGACAATTGGAGTGGAACCCAAACCGCTC  
 S G T T N T V A A Y N L T W K S T N F K T I L E W E P K P V 261 30

AATCAAGTCTACACTGTCATAAGCACTAACGACTAGCAGGAGATTGAAAAGCAATGCTTACACAACAGACACAGAGTGTGACCTCAC  
 N Q V Y T V Q I S T K S G D W K S K C F Y T T D T E C D L T 351 60

GACGAGATGTGAAGGATGTGAAGCAGACGTACTTGGCACGGCTTCTCTACCCGGCAGGGATGTGGAGAGCACCCGTTCTGGG  
 D E I V K D V K Q T Y L A R V F S Y P A G N V E S T G S A G 441 90

GAGCCTCTGTATGAGAACTCCCCAGAGTTCACACCTTACCTGGAGACAAACCTCGGACAGCCAACAATTAGCTTGAACAGGTGGG  
 E P L Y E N S P E F T P Y L E T N L G Q P T I Q S F E Q V G 531 120

ACAAAAGTGAATGTGACCGTAGAAGATGAACGGACTTACTGAGAACACACTTCTAACGCTCCGGATGTTGGCAAGGAC  
 T K V N V T V E D E R T L V R R N N T F L S L R D V F G K D 621 150

TTAATTTATACACTTATTATTGGAAATCTCAAGTTCAAGGAAAGACAGCCAAACAAACACTAATGAGTTTGATTGATGTGGAT  
 L I Y T L Y Y W K S S S S G K K T A K T N T N E F L I D V D 711 180

AAAGGAGAAAACACTGTTCAAGCAGTGATTCCCTCCGAACAGTTAACCGGAAGAGTACAGACAGCCGGTAGAGTGTATG  
 K G E N Y C F S V Q A V I P S R T V N R K S T D S P V E C M 801 210

GGCCAGGAGAAAGGGGAATTCAAGAGAAATATTCTACATCATGGAGCTGTGGCATTTGTGGTCATCATCTGTATCATCTGGCTATA  
 G Q E K G E F R E [I F Y I I G A V A F V V I I L V I I L A I 891 240

TCTCTACACAAGTGTAGAAAGCAGGAGTGGGCAGAGCTGGAAAGGAGAACCTCCACTGAATGTTCTAAAGGAAGCACTGTTGGAGC  
 S L J H K C R K A G V G Q S W K E N S P L N V S \* 981 219

TACTGCAAATGCTATATTGCACTGTGACCGAGAACCTTAAAGAGGATAGAACATGGAAACGCAAATGAGTATTCGGAGCATGAAGAC  
 CCTGGAGTCAAAAAACTCTTGATATGACCTGTTATTACCATAGCATTCTGGTTGACATCAGCATTAGTCACTTGAAATGTAACGA  
 ATGGTACTACAACCAATTCAAGTTAAATTAAACACCATGGCACCTTGCACATAACATGCTTAGATTATATTCCGCACTCAA  
 GGAGTAACCAAGGTGCTCCAAGCAAAACAAATGGAAAATGCTTAAAAAAATCTGGGTGGACTTTGAAAAGCTTTTTTTTTTT  
 TTTTGAGACGGAGTCTGCTCTGGCCAGGCTGGAGTGCAGTAGCATGATCTCGGCTACTGCACCCCTCCGCTCTGGGTTCAAGC  
 AATTGTCGCTCAGCCTCCGAGTAGCTGGATTACAGGTGGCACTACCAACACCAAGCTAATTGTTGATTAGAGATGGG  
 TTTCACCATCTGGCCAGGCTGGTCTGAATTCTGACCTCAGGTGATCCACCCACCTGGCTCCAAAGTGCTAGTATTATGGGCTG  
 AACCCACCATGCCAGCCAAAAGCTTTGAGGGCTGACTTCAATCCATGAGAACAGTAAATGGAAGGAAATTGGGTGCAATTCTAGG  
 ACTTTCTAACATATGCTATAATATAGTGTAGGTTCTTTTTTCAAGGAATACATTGAAATTCAAACATTGGCAAACATT  
 GTATTAAATGTTAAGTGCAGGAGACATTGGTATTCTGGGCACCTCTAAATATGCTTACATCTGCACTTAACTGACTTAAGTGGCA  
 TTAAACATTGAGAGCTAATATATTATAAGACTACTATACAAACTACAGAGTTATGTTAGGTACTAAAGCTTCTATGGTTG  
 ACATTGTTATATAATTGTTAAAAGGTTCTATATGGGATTCTATTACGTAGGTAAATTGTTCTATTGTTATATAATTGAGAT  
 AATTATTTAATATACTTTAAAGGTGGACTGGATTGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAA 2061 2151 2177

FIGURE 2: Nucleotide and translated amino acid sequence of human tissue factor cDNA isolates. Nucleotides and amino acids are numbered on the right. Nucleotide 1 was assigned to the first residue of the cDNA sequence, and amino acid 1 was assigned to the amino-terminal serine of mature tissue factor. The signal peptide is designated by negative numbers. The primary and proposed secondary signal peptidase cleavage sites are indicated by filled triangles. Potential N-linked glycosylation sites are marked by filled circles. Bold underlining indicates the amino-terminal protein sequence of tissue factor and also a potential polyadenylation or processing signal of AATAAA in the 3'-noncoding region. The proposed transmembrane domain is enclosed in brackets. The stop codon is indicated by the asterisk. The inverted Alu family repetitive sequence between flanking direct repeats is underlined. The positions of several intron-exon boundaries are indicated by arrows. The sequence represented in  $\lambda$ HTF8 includes nucleotides 21-1050, in  $\lambda$ HTF3 includes nucleotides 426-2078, in  $\lambda$ HTF7 includes nucleotides 1-~1250, and in  $\lambda$ HTF12 includes nucleotides 361-2177.

typical of signal sequences (von Heijne, 1983, 1985). The weighted matrix method of von Heijne (1983) predicts that the most likely site for signal peptidase to cleave is before Ser-1, and the second most likely site is before Thr-3. Thus, the amino-terminal heterogeneity observed in natural tissue factor may reflect the use of these two signal peptidase cleavage sites.

The sequence predicted for mature tissue factor consists of a probable extracellular domain of 219 amino acids that contains only four cysteine residues and only one methionine. This domain must contain a binding site for factor VII. The next 23 amino acids are extremely hydrophobic and probably

comprise a transmembrane segment that is flanked by charged residues. The carboxy-terminal 21 amino acids probably constitute a short cytoplasmic domain. The structural features of this domain resemble those of other transmembrane receptors such as the LDL receptor (Davis et al., 1987), including a cluster of positively charged residues adjacent to the plasma membrane, one cysteine residue, and several hydroxy amino acids. These features may be important for regulating the intracellular transport of tissue factor. The calculated amino acid composition agrees well with that reported for human brain tissue factor (Broze et al., 1985). The predicted mass of 29.6 kDa for mature tissue factor is considerably lower

Table I: Amino Acid Sequence of Human Brain Tissue Factor<sup>a</sup>

cycle	beginning with				Broze et al.	
	Ser-1		Thr-3			
	from cDNA	pmol detected	from cDNA	pmol detected		
1	S	≤40	S	T	17 T	
2	G	≤80	G	T	10 T X	
3	T	— X	N	20	N* N	
4	T	1.6 T	T	1.6	T T	
5	N	27 N*	V	28	V* V	
6	T	1.6 T	A	20	A* A	
7	V	23 V*	A	25	A* V	
8	A	22 A*	Y	14	Y* Y	
9	A	26 A*	N	2.6	N X	
10	Y	14 Y*	L	15	L* Y	
11	N	— X	T	1.6	T X	
12	L	10 L*	W	3.5	W L	
13	T	1.6 T	K	9.8	K* K	
14	W	4.2 W	S	—	X (S)	
15	K	7.4 K*	T	—	X K	
16	S	— X	N	5.1	N* N	
17	T	— X	F	5.3	F* F	
18	N	— X	K	2.9	K*	
19	F	4.4 F*				

<sup>a</sup> Human brain tissue factor was purified and sequenced as described previously (Broze et al., 1985) with ~190 pmol of protein. The amino acid sequence predicted from the cDNA sequence is aligned with the residues detected at each cycle of automated Edman degradation for two major sequences and compared to the sequence reported by Broze et al. (1985) in the last column. Residues marked by asterisks were used to estimate the initial and repetitive yields by linear regression analysis as described in the text.

than the 44 kDa estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Broze et al., 1985), and this difference probably reflects in part the contribution of glycosylation to the mobility of the natural protein. Removal of N-linked carbohydrate from brain or placenta tissue factor has been reported to reduce the apparent mass by ~10 kDa (Morrissey et al., 1986). The extracellular domain of tissue factor contains three potential glycosylation sites with the sequence Asn-X-Thr/Ser (Figure 2). Residue Asn-11 was not clearly detected by protein sequencing (Table I), suggesting that it is glycosylated.

Morrissey et al. (1987a) have recently claimed that tissue factor consists of a 47-kDa glycosylated heavy chain disulfide-linked to a 12.5-kDa light chain. The data presented here indicate that tissue factor is probably synthesized as a single polypeptide precursor with a typical cleavable signal peptide before the amino terminus that is found in the tissue factor protein purified from human brain. Therefore, a light chain component, if present, would probably be derived from a separate gene and may not be necessary for cofactor function (Broze et al., 1985).

**Comparison of Tissue Factor to Other Proteins.** Searches of the NBRF Protein Sequence Database and the Genbank Genetic Sequence Data Bank did not find any previously characterized proteins that might be homologous to tissue factor. Three of the five tryptophan residues of mature tissue factor occur in the tripeptide sequence Trp-Lys-Ser. The tripeptide sequence does occur in several other proteins, and the physiological significance of this observation is not known. Aside from this structural feature there is no evidence for internal sequence duplication in human tissue factor.

**Size and Occurrence of Tissue Factor mRNA in Tissues and Cultured Cells.** A major tissue factor mRNA of 2.2 kb was detected in brain and placenta (Figure 3). The agreement between the length of the mRNA and the length of the composite cDNA sequence (Figure 2) suggests that the cDNA

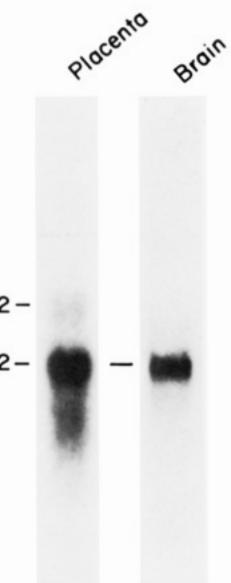


FIGURE 3: Northern blot of RNA from brain and placenta probed with the tissue factor cDNA insert of  $\lambda$ HTF8. The lanes contain 10  $\mu$ g of placenta or brain poly(A)+ RNA. The sizes of tissue factor mRNA species were estimated by interpolation between the positions of  $\lambda$ -phage *Hind*III fragment markers and are indicated at the left.

sequence is essentially full length. A diffuse minor band of ~3.2 kb was found in placenta. The relationship between the major and minor species is not yet known. The ~3.2-kb species may arise from either incomplete or alternative RNA splicing in placenta. Several additional cDNA isolates from placenta have been characterized that appear to contain unprocessed introns with typical splice junctions and these may correspond to the ~3.2-kb RNA species. Partial sequences determined for these isolates suggest that the gene will contain intron-exon boundaries between nucleotides 174-175, 666-667, and 826-827 (unpublished data). There is some correlation between the placement of these splice junctions and the domains of the protein. The first occurs at the boundary between the signal peptide and the mature protein sequence, and the latter occurs at the extracellular boundary of the transmembrane segment (Figure 2). Whether these transcripts give rise to different functional tissue factor protein sequences is not known. Only the 2.2-kb tissue factor mRNA was found in the human monocytic U937 cell line, and the concentration of this species was increased severalfold by treatment of the cells for 12 h with 1  $\mu$ M phorbol 12-myristate 13-acetate. Control hybridizations with a  $\gamma$ -actin cDNA probe (Gunning et al., 1983) showed that comparable amounts of RNA had been transferred (data not shown). This parallels the previously reported induction of tissue factor activity in these cells by phorbol esters (Lyberg et al., 1982).

**Chromosome Localization of the Human Tissue Factor Gene.** The cDNA insert of  $\lambda$ HTF8 was hybridized to human chromosomes purified by fluorescence-activated flow sorting. Two complete sets of 22 autosomes and the X and Y chromosomes were examined, both of which gave signals only with chromosome 1. A selection of these filters is shown in Figure 4. This assignment agrees with that made earlier by the determination of tissue factor activity in mouse-human somatic cell hybrid cell lines with defined complements of human chromosomes (Carson et al., 1985).

The cloning of an essentially full-length cDNA for human tissue factor completes the determination of the primary structure for all known components of the classical intrinsic and extrinsic blood clotting cascades. These clones provide a foundation for future investigations of the biosynthetic

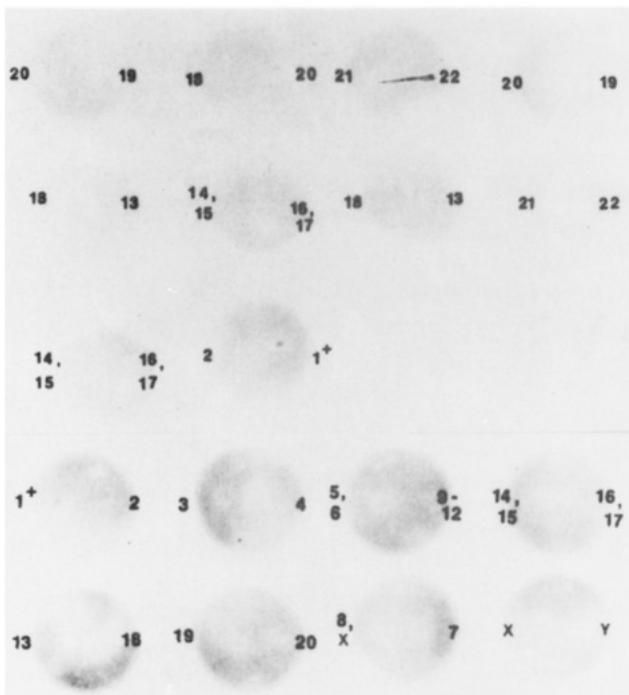


FIGURE 4: Chromosome localization of the human tissue factor gene. The characters (1-22, X, Y) indicate the human chromosomes present in each spot. The positive signals obtained with DNA from chromosome 1 are marked (+).

regulation and structure-function relationships of tissue factor.

#### ADDED IN PROOF

After acceptance of this manuscript, a partial cDNA sequence for human tissue factor containing the entire coding sequence was published by Morrissey et al. (1987b). The two reports agree except that their nucleotides 1-4 are not found in any of our three independent cDNA isolates for that region, and for amino acid residue 228 we find alanine (GCA) instead of valine (GTA). This latter discrepancy is probably due to a sequence polymorphism.

#### ACKNOWLEDGMENTS

We thank Rosalind Kornfeld (Washington University) for encouragement and helpful discussions, Karen O'Malley (Washington University) for providing human brain poly(A)+ RNA, Laurence Kedes (Stanford University, Stanford, CA) for the  $\gamma$ -actin cDNA, Richard Ye (Washington University) for providing RNA from U937 cells, Beverly Shelton-Inloes (Washington University) for help with DNA sequencing, Lisa Westfield (Washington University) for the synthesis of oligonucleotides, and Roger V. Lebo (University of California, San Francisco) for the sorting of human chromosomes.

#### REFERENCES

Bach, R., Nemerson, Y., & Konigsberg, W. (1981) *J. Biol. Chem.* 256, 8324-8331.

Brox, J. H., Østerud, B., Bjørklid, E., & Fenton, J. W., II (1984) *Br. J. Haematol.* 57, 239-246.

Broze, G. J., Jr., Leykam, J. E., Schwartz, B. D., & Miletich, J. P. (1985) *J. Biol. Chem.* 260, 10917-10920.

Caput, D., Buetler, B., Hartog, K., Thayer, R., Brown-Shimer, S., & Cerami, A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 1670-1674.

Carson, S. D., Henry, W. M., & Shows, T. B. (1985) *Science (Washington, D.C.)* 229, 991-993.

Davis, C. G., van Driel, I. R., Russell, D. W., Brown, M. S., & Goldstein, J. L. (1987) *J. Biol. Chem.* 262, 4075-4082.

Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T., & Schmid, C. W. (1981) *J. Mol. Biol.* 151, 17-33.

Guha, A., Bach, R., Konigsberg, W., & Nemerson, Y. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 299-302.

Gunning, P., Ponte, P., Okayama, H., Engel, J., Blau, H., & Kedes, L. (1983) *Mol. Cell. Biol.* 3, 787-795.

Kozak, M. (1986) *Cell (Cambridge, Mass.)* 44, 283-292.

Lebo, R. V., Gorin, F., Fletterick, R. J., Kao, F. T., Cheung, M. C., Bruce, B. D., & Kan, Y. W. (1984) *Science (Washington, D.C.)* 225, 57-59.

Lyberg, T., Nilsson, K., & Prydz, H. (1982) *Br. J. Haematol.* 51, 631-641.

Lyberg, T., Galdal, K. S., Evensen, S. A., & Prydz, H. (1983) *Br. J. Haematol.* 53, 85-95.

Morrissey, J. H., Fair, D. S., Guha, A., Bach, R., Nemerson, Y., & Edgington, T. S. (1986) *Circulation* 74, II-409.

Morrissey, J. H., Fakhrai, H., & Edgington, T. S. (1987a) *Fed. Proc., Fed. Am. Soc. Exp. Biol.* 46, 716.

Morrissey, J. H., Fakhrai, H., & Edgington, T. S. (1987b) *Cell (Cambridge, Mass.)* 50, 129-135.

Nawroth, P. P., & Stern, D. M. (1986) *J. Exp. Med.* 163, 740-745.

Nawroth, P. P., Handley, D. A., Esmon, C. T., & Stern, D. M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 3460-3464.

Østerud, B., & Rapaport, S. I. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5260-5264.

Østerud, B., & Bjørklid, E. (1982) *Scand. J. Haematol.* 29, 175-184.

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

Rivers, R. P. A., Hathaway, W. E., & Weston, W. L. (1975) *Br. J. Haematol.* 30, 311-316.

Scarpatti, E. M., Broze, G. J., Miletich, J. P., Lebo, R. V., & Sadler, J. E. (1987) *Fed. Proc., Fed. Am. Soc. Exp. Biol.* 46, 2242.

Shaw, G., & Kamen, R. (1986) *Cell (Cambridge, Mass.)* 46, 659-667.

Sigel, M. B., Sinha, Y. N., & VanderLaan, W. P. (1983) *Methods Enzymol.* 93, 3-12.

Silverberg, S. A., Nemerson, Y., & Zur, M. (1977) *J. Biol. Chem.* 252, 8481-8488.

Stern, D. M., Drillings, M., Kisiel, W., Nawroth, P., Nossel, H. L., & LaGamma, K. S. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 913-917.

Straub, W., & Duckert, F. (1961) *Thromb. Diath. Haemorrh.* 5, 402-404.

Tanaka, H., Janssen, B., Preissner, K. T., & Müller-Berghaus, G. (1985) *Thromb. Res.* 40, 745-756.

Vaitukaitus, J. L. (1981) *Methods Enzymol.* 73, 46-52.

von Heijne, G. (1983) *Eur. J. Biochem.* 133, 17-21.

von Heijne, G. (1985) *J. Mol. Biol.* 184, 99-105.

Weinberger, C., Hollenberg, S. M., Ong, E. S., Harmon, J. M., Brower, S. T., Cidlowski, J., Thompson, E. B., Rosenfeld, M. G., & Evans, R. M. (1985) *Science (Washington, D.C.)* 228, 740-742.

Wen, D., Dittman, W. A., Ye, R. D., Deaven, L. L., Majerus, P. W., & Sadler, J. E. (1987) *Biochemistry* 26, 4350-4357.

Williams, W. J. (1964) *J. Biol. Chem.* 239, 933-942.

Williams, W. J. (1966) *J. Biol. Chem.* 241, 1840-1846.

Ye, R. D., Wun, T.-C., & Sadler, J. E. (1987) *J. Biol. Chem.* 262, 3718-3725.