D. catawba × *schoedleri* after J. L. Brooks, "The systematics of North American *Daphnia*," *Mem. Conn. Acad. Arts Sci.* **12** (1957); individuals in the population show features of species within the *Pulex* group but come closest to resembling *D. catawba*.
24. Detailed studies in nearby Lake 114 show a pronounced overlap in the diets of *S. margarita* and *P. promelas* [R. F. Tallman, K. H. Mills, R. G. Rotter, *Can. Manuscr. Rep. Fish. Aquat. Sci.* (No. 1756) (1984)].
25. R. L. France and L. Graham, *Water Air Soil Sci.*, in press; R. L. France, thesis, University of Manitoba, Winnipeg (1982); crayfish populations were estimated by sampling (SCUBA) known areas along transects at night. Numbers of leeches, sculpins, *Hexagenia*, and tadpoles seen were also recorded (I. J. Davies, unpublished data).
26. Spawning was observed each fall from 1979 to 1983 by K. H. Mills by patrolling the shoreline at night with a light-equipped boat.
27. Similar species compositions in acidified lakes have been reported by N. D. Yan and P. M. Stokes, *Environ. Conserv.* **5**, 93 (1982); B. Almer, W. Dickson, C. Eckstrom, E. Hornstrom, *Ambio* **3**, 30 (1974); E. Hornstrom, C. Eckstrom, U. Miller, W. Dickson, *Information Fran Sotvattens Laboratoriet, Drottningholm* (No. 4) (1973).
28. D. W. Schindler, R. Wagemann, B. deMarch, M. A. Turner, in preparation.
29. D. W. Schindler and M. A. Turner, *Water Air Soil Pollut.* **18**, 259 (1982).
30. E. J. Fee, D. Hayward, J. A. Shearer, *Can. Data Rep. Fish. Aquat. Sci.* (1982), p. 327; J. A. Shearer and E. R. DeBruyn, unpublished data.
31. M. A. Turner, unpublished data.
32. O. Grahn, H. Hultberg, L. Landner, *Ambio* **3**, 8 (1974); W. Dickson, "Properties of acidified waters," in (*8*), p. 75.
33. Natural lakes in the ELA are phosphorus-limited. The slight increases in phytoplankton production and biomass observed in Lake 223 appeared to be due to low water renewal of the lake caused by several years of lower than normal precipitation in the period of study; D. W. Schindler, *Science* **195**, 260 (1977); ———, E. J. Fee, T. Ruszczynski, *J. Fish. Res. Board Can.* **35**, 190 (1978).
34. C. A. Kelly, J. W. M. Rudd, A. Furutani, D. W. Schindler, *Limnol. Oceanogr.* **29**, 687 (1984).
35. J. Økland and K. Økland, in (*8*), p. 326; L. Henrikson, H. G. Oscarson, J. A. E. Stenson, in (*8*), p. 316; J. P. Nilssen, in (*8*), p. 344.
36. D. I. Mount, *Water Res.* **7**, 987 (1973).
37. See review by T. A. Haines, *Trans. Am. Fish. Soc.* **110**, 669 (1981) and papers in "Acid Rain/Fisheries," Proceedings of an international symposium on acid rain and fishery impacts on northeastern North America (Cornell University, Ithaca, New York, 1981).
38. A. Herczeg and R. H. Hesslein, *Geochim. Cosmochim. Acta* **48**, 837 (1984).
39. For example, T. S. Traaen, "Effects of acidity on decomposition of organic matter in aquatic environments," in (*8*), p. 340. P. Ramlal and J. W. M. Rudd, in preparation; G. R. Hendrey, "Effects of *p*H on the growth of periphytic algae in artificial stream channels" (Report IR 25/76, SNSF Project, Oslo, Norway, 1976).
40. D. C. Lasenby and R. R. Langford, *J. Fish. Res. Board Can.* **29**, 1701 (1972); *Limnol. Oceanogr.* **18**, 280 (1973).
41. We have previously shown that by not incorporating atmospheric exchange and natural succession of algal species, laboratory experiments led to erroneous conclusions regarding the control of eutrophication; see D. W. Schindler in (*33*).
42. G. Davidson, thesis, University of Manitoba (1984).
43. In order to compare relative temporal variation in Lake 223 and reference lakes in panels B, C, D, and J of Fig. 1, the following procedure was used. (i) The long-term mean concentration or rate for a variable in each lake was calculated for the period 1974 to 1982. The time-weighted average annual mean for each open water year was calculated from weekly to monthly samples and then divided by the long-term mean, resulting in a group of numbers clustered about 1.0. These clusters did not depart significantly from normal distributions. For each year, the average ratio of the annual mean to the long-term mean was calculated for all control lakes, and 95 percent confidence limits were calculated. For the four examples shown, data from two to eight control lakes were used in any year, varying somewhat with the difficulty and cost of measurement. Lake 223 annual averages, normalized to the long-term mean as described above are shown for comparison with reference lakes. The following data allow conditions in lake 223 before acidification to be compared to long-term means for four to six reference lakes in the area, for the period 1974 to 1982. Means and standard deviations are given for reference lakes, followed by the preacidification value for Lake 223 in brackets. Phytoplankton production, as carbon, 164 ± 60 mg/m^2 per day (140 ± 56); phytoplankton biomass, 1.03 ± 0.41 mg/liter (0.69 ± 12); chlorophyll a, 3.00 ± 0.84 μg/liter (1.92 ± 0.14); dissolved inorganic carbon, 1.41 ± 0.61 mg/liter (1.60 ± 0.07); Ca, 2.43 ± 0.85 mg/liter (2.19 ± 0.02); SO_4, 4.09 ± 0.99 mg/liter (3.35 ± 0.20); Si, 1.20 ± 1.04 mg/liter (1.29 ± 0.01); total nitrogen, 314 ± 51 μg/liter (272 ± 25); and total phosphorus, 6.6 ± 1.2 μg/liter (6.8 ± 0.1). Mean depth and lake surface area hectare for Lake 223 and the reference lakes (mean and standard deviation, n = 6) were 7.2 m versus 7.2 m ± 3.3 m, and 27.3 ha versus 31.3 ha ± 16.6 ha, respectively.
44. E. H. Simpson, *Nature (London)* **163**, 688 (1949).
45. R. Playle did many of the calculations for Fig. 1. P. Campbell, D. Rosenberg, R. Hecky, E. deNoyelles, Jr., G. D. Koshinsky, and K. Patalas provided helpful reviews of this manuscript. This work was supported by the Canadian Department of Fisheries and Oceans and the Alberta Oil Sands Environmental Research Project (AOSERP).

5 November 1984; accepted 2 May 1985

RESEARCH ARTICLE

Human von Willebrand Factor (vWF): Isolation of Complementary DNA (cDNA) Clones and Chromosomal Localization

David Ginsburg, Robert I. Handin, David T. Bonthron
Timothy A. Donlon, Gail A. P. Bruns
Samuel A. Latt, Stuart H. Orkin

The hemostatic system has evolved to minimize blood loss following vascular injury. In higher vertebrates, including man, the system is quite complex and requires the interaction of circulating platelets, a series of plasma coagulation proteins, endothelial cells, and components of the vascular subendothelium. The initial and critical event in hemostasis is the adhesion of platelets to the subendothelium. It occurs within seconds of injury and provides a nidus for platelet plug assembly and fibrin clot formation.

The factor VIII molecular complex is composed of two distinct protein components, the antihemophilic factor (AHF or VIIIC) and von Willebrand factor (vWF), and plays a major role in both platelet adhesion and fibrin formation (*1*). Two of the most common inherited clinical bleeding disorders are the result of a deficiency in the activity of one or the other of these components.

The VIIIC molecule is an important regulatory protein in the coagulation cascade. After activation by trace quantities of thrombin, it accelerates the rate of factor X activation by factor IX, eventually leading to the formation of the fibrin clot. Classic hemophilia (VIIIC deficiency) is an X chromosome–linked disorder that affects one in 10,000 males, and has been recognized as a major source of hemorrhagic morbidity and mortality since biblical times. Treatment consists of supportive measures and usually requires frequent transfusion with blood products. The latter results in a high incidence of infectious complications in this population, including various forms of hepatitis and acquired immune deficiency syndrome.

The vWF molecule is an adhesive glycoprotein synthesized by endothelial cells and megakaryocytes. It serves as a carrier in plasma for VIIIC and facilitates platelet-vessel wall interactions. By binding to subendothelial structures and to the platelet surface it promotes shear-dependent platelet adhesion to the vessel

The authors are members of the Departments of Medicine, Pediatrics, and Genetics at Harvard Medical School, Boston, Massachusetts. David Ginsburg and Robert Handin are also members of the Hematology Division, Department of Medicine, Brigham and Women's Hospital, Boston. Stuart Orkin and David Bonthron are members of the Hematology Division, Children's Hospital and Dana-Farber Cancer Institute, Boston. Tim Donlon, Gail Bruns, and Sam Latt are members of the Genetics Division, Children's Hospital, Boston.

wall. Discrete domains of vWF which bind to platelet receptor sites on glycoprotein Ib and on the glycoprotein IIb–IIIa complex, as well as to binding sites on collagen, have been noted. A variety of abnormalities in vWF activity can result in von Willebrand's disease (vWD). This disorder is generally inherited in an autosomal dominant fashion and may affect as many as one in 2000 individuals. The mild forms may frequently go undiagnosed. Severely affected patients may require frequent blood product support with its associated risks.

Abstract. *Human factor VIII–von Willebrand factor (vWF) is a large, multimeric glycoprotein that plays a central role in the blood coagulation system, serving both as a carrier for factor VIIIC (antihemophilic factor) and as a major mediator of platelet–vessel wall interaction. Diminished or abnormal vWF activity results in von Willebrand's disease (vWD), a common and complex hereditary bleeding disorder. Overlapping vWF cDNA clones that span 8.2 kilobases of the vWF messenger RNA have been obtained. vWF accounts for approximately 0.3 percent of endothelial cell messenger RNA and was undetectable in several other tissues examined. A large single copy gene for vWF is located on the short arm of chromosome 12 (12p12 → 12pter). No gross gene rearrangement or deletion was detected in the DNA of two patients with severe vWD.*

Whereas the VIIIC molecule is a single-chain 220-kilodalton (kD) protein, vWF activity is expressed in a heterogeneous series of multimers with molecular sizes ranging from 450 to 20,000 kD. These multimers are assembled from a single glycoprotein subunit of approximately 220 kD. vWF accounts for 99 percent of the mass of the plasma factor VIII molecular complex (*1*). The VIIIC messenger RNA (mRNA) has been found in various tissues including liver, placenta, and a T-cell hybridoma line (*2*). The vWF protein has only been detected in endothelial cells, megakaryocytes, and in tumors derived from vascular tissue (*1*, *3*). In cultured endothelial cells, vWF is synthesized as a large precursor (240 to 260 kD) which is processed to the mature 220-kD subunit and assembled into multimers which then enter plasma (*4*). Endothelial cells are the major site of plasma vWF synthesis. Blood platelet alpha granules contain approximately 15 percent of circulating vWF. There is little information regarding megakaryocyte biosynthesis of vWF or the role of this platelet vWF pool in hemostasis.

As noted above, vWD is a complex and heterogeneous group of hereditary bleeding disorders (*1*). In the most common variety (type I), patients have reduced vWF activity, but secrete the full range of multimers into plasma. Several additional forms of vWD have been defined from clinical studies. Although their molecular defects are not fully understood, they all exhibit a qualitative or quantitative abnormality of the vWF molecule. The most common variants have a selective loss of the high molecular weight multimers of vWF due to a failure of assembly (type IIa) or rapid clearance due to aberrant platelet binding (type IIb). Occasional patients with an autosomal recessive form of vWD have been described. Analyses of several large kindreds have failed to demonstrate definite linkage of any form of vWD to other genetic markers (*5*).

The study of vWF biochemistry, structure, and activity has been particularly difficult because of its large size, heterogeneous nature, and poorly defined function. Until recently, there was no information on the primary structure of either the VIIIC or vWF proteins, or on the molecular basis of their respective clinical defects. The isolation of genomic and complementary DNA (cDNA) clones for VIIIC has now been reported (*2*), and molecular techniques should soon be applicable to the prenatal diagnosis of hemophilia A. We have isolated cloned vWF cDNA sequences, which will now permit a molecular genetic approach to the study of vWD and to the analysis of vWF structure and function.

Expression cloning of vWF sequences. In order to isolate vWF cDNA we used the λgt11 bacteriophage expression vector with a specific antibody as the detection system. In this expression system, devised by Young and Davis (*6*), proteins are produced as fusion products at the COOH-terminus of bacterial β-galactosidase. A number of cDNA's have been cloned in this way (*6–7*), including those for several coagulation proteins (*8*).

We constructed two large λgt11 bacteriophage libraries of cDNA derived from cultured human umbilical vein endothelial cells (HUVEC), which are a recognized source of vWF. For the first library cDNA was synthesized with oligodeoxythymidylate [oligo(dT)] as primer [oligo(dT) library], and a pool of random hexanucleotides was used as primer for the second (random library). The random library was expected to contain clones randomly distributed along the length of the mRNA template, a feature that would avoid the bias toward the 3′ end inherent in libraries primed with oligo(dT).

A primary culture of HUVEC (*9*) was grown and passaged in Medium 199 with 20 percent fetal bovine serum in the presence of bovine endothelial cell growth factor (*10*) and fibronectin by the method of Maciag *et al.* (*11*). Growth was markedly enhanced by addition of heparin as described by Thornton *et al.* (*12*). Cultured cells were positive for

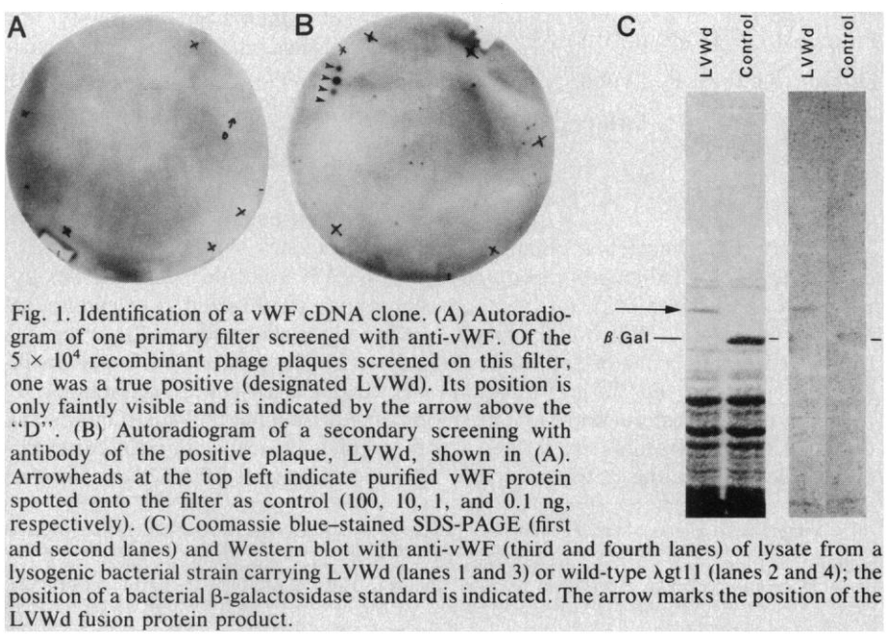
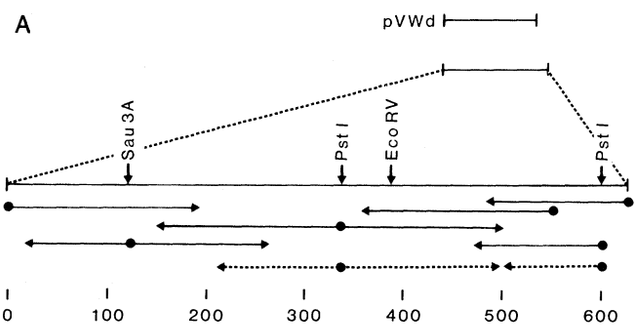


Fig. 1. Identification of a vWF cDNA clone. (A) Autoradiogram of one primary filter screened with anti-vWF. Of the 5×10^4 recombinant phage plaques screened on this filter, one was a true positive (designated LVWd). Its position is only faintly visible and is indicated by the arrow above the "D". (B) Autoradiogram of a secondary screening with antibody of the positive plaque, LVWd, shown in (A). Arrowheads at the top left indicate purified vWF protein spotted onto the filter as control (100, 10, 1, and 0.1 ng, respectively). (C) Coomassie blue–stained SDS-PAGE (first and second lanes) and Western blot with anti-vWF (third and fourth lanes) of lysate from a lysogenic bacterial strain carrying LVWd (lanes 1 and 3) or wild-type λgt11 (lanes 2 and 4); the position of a bacterial β-galactosidase standard is indicated. The arrow marks the position of the LVWd fusion protein product.

vWF antigen by immunofluorescence (*11*), and the conditioned media from the cultures contained vWF antigen as determined by ELISA assay. After four additional passages, cells were harvested and total RNA prepared by guanidine-HCl extraction (*13*). Polyadenylated [poly(A)+] mRNA was isolated from total endothelial cell RNA by oligo(dT)-cellulose column chromatography (*14*). cDNA was synthesized by the method of Okayama and Berg (*15*) as modified by Gubler and Hoffman (*16*). As noted above, one library was prepared with oligo$(dT)_{12-18}$ as primer for the first strand synthesis [oligo (dT) library] while a second cDNA pool was synthesized with a random mixture of hexanucleotides (P-L Biochemicals) as primer for first strand synthesis (random library). T4 DNA polymerase was used to create blunt ends and the cDNA's were ligated to synthetic Eco RI linkers after protection of internal Eco RI sites. The linker-ligated cDNA's were digested with Eco RI and separated from free linkers by Sepharose CL4B chromatography. Complementary DNA's were then ligated to Eco RI–digested, alkaline phosphatase–treated λgt11 vector DNA (*6*), packaged in vitro (*14*), and plated on *Escherichia coli* host strain Y1088 (*6*). Each HUVEC cDNA library contained approximately 3×10^6 to 4×10^6 independent recombinant clones. Nonrecombinant background, as assessed by growth on isopropyl thio-β-D-galactopyranoside (IPTG)–X-Gal plates (*6*), was approximately 30 percent. Inserts of ten randomly chosen cDNA clones were 1 to 3 kb in length.

For detection of bacterial clones harboring recombinant phage encoding products of the fusion of vWF and β-galactosidase, we used an affinity purified heteroantiserum to human vWF (anti-vWF) (*17*). Approximately 3×10^6 recombinant clones from the above λgt11 endothelial cell cDNA libraries were screened as phage plaques on *E. coli* host strain Y1090 with anti-vWF at a 1:1000 dilution, by the method of Young and Davis (*6*) (Fig. 1A). One of nine candidate positive plaques was strongly positive on a repeat antibody screening (Fig. 1B). Phage DNA of this clone (LVWd) was prepared by standard methods (*14*). Lysogens of LVWd and wild-type λgt11 were prepared in the *E. coli* bacterial host strain Y1089 (*6*). Lysogens were grown in LB medium at 32°C to $OD_{550} = 0.5$. Phage were induced by temperature shift to 45°C for 20 minutes, and β-galactosidase production was induced by the addition of IPTG (2 to 5 m*M* final concentration). After an additional 1 hour at 38°C, bacteria were harvested and resuspended in TBS with 0.2 m*M* phenylmethylsulfenyl fluoride. Lysates were prepared by two freeze-thaw cycles and sonication, and examined by gel electrophoresis and "Western blot" analysis (*17*). The β-galactosidase fusion protein material synthesized by the LVWd lysogen stained specifically with anti-vWF (Fig. 1C).

LVWd is authentic vWF cDNA. The nucleotide sequence was determined (Fig. 2) for the 549-base-pair cDNA insert of LVWd and an additional 81 bp just 3′ to it, which was obtained from the overlapping clone pVWE6 (Fig. 3). A single open reading frame encoding 193 amino acids was followed by a single termination codon. In LVWd this sequence was in the same orientation and reading frame as the β-galactosidase gene with which it was fused. This is consistent with expression of a fusion protein product. To provide evidence independent of immunodetection that this cDNA encodes vWF, the primary COOH-terminal amino acid sequence was determined by limited carboxypeptidase Y digestion of purified vWF protein and subsequent high-performance liquid chromatography (HPLC) analysis of cleaved amino acids. The sequence obtained (K-C-S-K) (*18*) was identical to that immediately preceding the stop codon in our cDNA (Fig. 2B). In addition, the predicted COOH-terminal amino acid sequence E-C-K-C-S-P-R-K-C-S-K



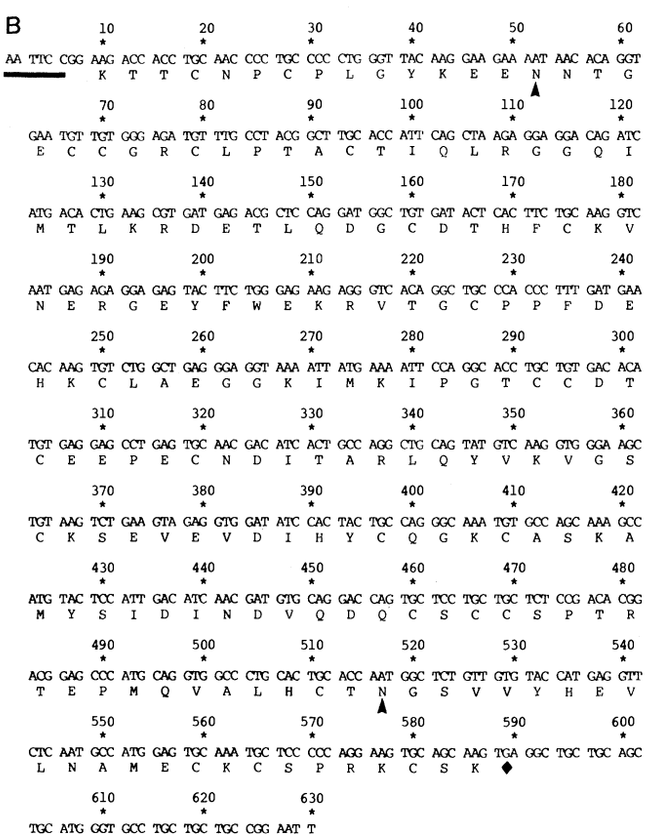


Fig. 2. (A) Restriction map of the insert of clone LVWd with the 3′ end of the overlapping clone pVWE6 (see Fig. 3), and the sequencing strategy employed. Solid lines and arrows indicate regions sequenced by the method of Maxam and Gilbert (*30*) with solid circles indicating the end-labeled restriction site. Dotted lines and arrows indicate regions sequenced by the method of Sanger *et al.* (*31*) with PstI fragments subcloned into M13 mp11 (*32*). (B) DNA sequence for 618 bp at the 3′ end of vWF cDNA. The predicted amino acid sequence for 193 amino acid residues at the COOH-terminus of vWF is shown in the single letter amino acid code (*18*). The six nucleotides at the beginning and end of the DNA sequence correspond to the synthetic Eco RI linker introduced by the cloning procedure. The single termination codon is marked by a diamond. Arrowheads indicate two potential *N*-glycosylation sites.

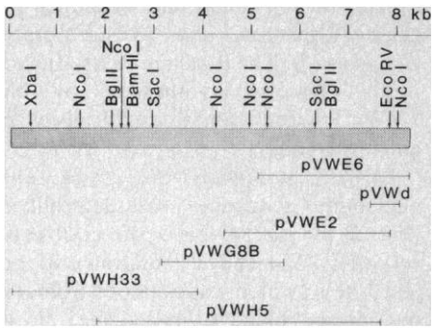


Fig. 3. Restriction map of vWF cDNA. Overlapping cDNA clones obtained by rescreening of the HUVEC library were used to construct this map. The locations of the original clone pVWd and selected overlapping clones are shown. Of the six vWF cDNA clones illustrated here, only pVWH33 was isolated from the random library. Others were obtained from the oligo(dT) library. Complementary DNA inserts of LVWd and other recombinant phage isolates were subcloned (for example, pVWd) into Eco RI–cut and bovine alkaline phosphatase–treated pUC-13 plasmid (P-L Biochemicals). The probes and strategy used are described in the text. Fragments were ^{32}P-labeled by an oligonucleotide primer labeling method (*33*).

(Fig. 2B) exactly matches the 11-residue COOH-terminal primary amino acid sequence reported by Titani *et al.* (*19*). We conclude that LVWd encodes vWF and not another protein species inadvertently detected by antibody screening.

Overlapping cDNA clones spanning the vWF transcript. The insert of LVWd was used as probe to rescreen the HUVEC libraries. Positive phage were detected with an abundance of about 0.3 percent. Seven of the vWF recombinant phage were plaque-purified and the four largest cDNA inserts were subcloned into the plasmid vector pUC-13. A 270-bp Eco RI–Pst I fragment from the 5′ end of pVWE2 (see Fig. 3) was used as probe to rescreen both HUVEC cDNA libraries, and a third series of overlapping vWF cDNA clones was identified. A 400-bp Stu I fragment from the 5′ end of one of these clones (pVWG8b) was then used as probe for a fourth round of screening of the primary libraries, and another set of cDNA clones further toward the 5′ end was obtained. These overlapping clones span 8.2 kb of vWF mRNA (Fig. 3).

The nucleotide sequence shown in Fig. 2B contains the translation termination signal TGA (*18*) at a position corresponding to 8.15 kb in the vWF cDNA map (Fig. 3). The remaining 5′ potential coding sequence is sufficient in length to encode a polypeptide of 300 kD, assuming a mean residue molecular weight of 111 based on the amino acid composition of vWF (*20*). The termination codon is followed by 34 bases of 3′ untranslated region (Fig. 2B). Neither the conserved endonucleolytic poly(A) addition signal AATAAA (*21*) nor a 3′ poly(A) tract are contained in this sequence. We surmise that additional 3′ untranslated sequences exist in vWF mRNA.

As is discussed above, the COOH-terminal sequence obtained by primary amino acid sequence analysis of purified vWF polypeptide is identical to the predicted amino acid sequence immediately preceding the stop codon in our cDNA (Fig. 2B). These data indicate that no posttranslational processing occurs at the COOH-terminal end of the vWF polypeptide.

vWF protein structure. Previous biochemical studies have demonstrated that vWF is rich in proline (6.8 percent) and cysteine (7.3 percent) (*20*) with each subunit participating in 69 intrachain and interchain disulfide bonds (*20, 22*). The predicted sequence of vWF shown in Fig. 2B contains 5.2 percent proline and 12.4 percent cysteine in 193 residues. This COOH-terminal fragment also contains two potential *N*-glycosylation sites (Fig. 2B).

Expression of vWF mRNA. When used as a hybridization probe in Northern blot analysis (Fig. 4), vWF cDNA detects a single large (approximately 8 to 10 kb) endothelial cell–specific RNA species. The observed signal intensity is in accord with the 0.3 percent frequency of positive phage detected on rescreening of the HUVEC cDNA library. No hybridization was observed with RNA's from human fibroblasts, HeLa cells, a human T-cell leukemia line (Fig. 4), or human kidney (*23*).

Gene structure and chromosomal localization. Portions of the vWF cDNA have been used as probes in Southern blot analyses to assess the complexity of the cellular vWF gene and its chromosomal localization. Initial use of LVWd as a probe revealed a simple pattern of hybridization, consistent with the pres-

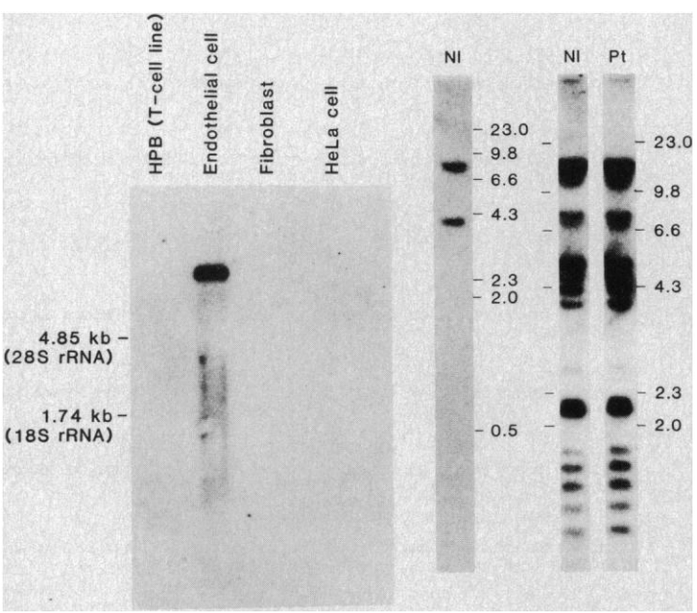


Fig. 4 (left). Northern blot analysis. Total cell RNA from HPB-ALL (a T-cell line) (lane 1), endothelial cells (HUVEC) (lane 2), fibroblasts (lane 3), and HeLa cells (lane 4) were hybridized with pVWd cDNA insert as probe. The location of ribosomal RNA size markers is indicated. A single hybridizing band of approximately 8 to 10 kb was evident in the endothelial cell lane. No signal was detectable in the other lanes. Total RNA was prepared from endothelial cells (HUVEC) and nonendothelial cells (HeLa cells, fibroblasts, and the T-cell line HPB-ALL), by lysis in a buffer containing 6*M* guanidine-HCl (*13*). The RNA's (20 μg of total RNA per lane) were denatured with formaldehyde, subjected to electrophoresis in agarose, and transferred to nitrocellulose filters (*14*). The filters were hybridized with ^{32}P-labeled LVWd insert. Hybridization was in 10 percent dextran, 50 percent formamide, 5× standard saline citrate (SSC) (0.75*M* NaCl, 0.075*M* sodium citrate) at 42°C; the final washing was done in 0.1× SSC at 68°C. Autoradiography with an intensifying screen at −80°C was for 24 hours (above), or 10 days (not shown). Fig. 5 (right). Southern blot analysis (*14*) of normal and vWD patient genomic DNA. Lane 1 shows normal human DNA digested with Bam HI with LVWd insert as probe. DNA's in lanes 2 and 3 have been digested with Eco RI and probed with nearly complete vWF cDNA (that is, pooled radiolabeled inserts of pVWH5 and pVWH33; see Fig. 3). The DNA in lane 2 was prepared from a normal individual, and the DNA in lane 3 from a patient with severe vWD. Size marker positions are indicated to the right of lane 1 and lanes 2 to 3. Hybridization and washing conditions were identical to those described in Fig. 4 above, except that dextran was omitted.

ence of a single copy in the human genome (Fig. 5). Probes spanning almost the complete vWF mRNA detected a complex array of 17 or more bands (Fig. 5). This pattern indicates a large gene interrupted by at least 16 introns and spanning a minimum of 80 kb of human genome. No rearrangement or deletion of vWF sequences was detected in DNA obtained from a patient with a severe variant of vWD (Fig. 5). The patient had no detectable vWF antigen (VIIIR:Ag, 0 percent); this is measured with the use of an antibody to the human factor VIII-vWF complex, with pooled normal plasma as control. Ristocetin cofactor activity (VIIIR:RCo) was 10 percent of control. This is a functional assay of vWF based on the observation that the antibiotic ristocetin induces platelet aggregation in the presence of vWF (*1*). Analysis of DNA from a second patient with severe vWD (VIIIR:Ag, 0 percent, VIIIR:RCo, 0 percent), similarly yielded a pattern indistinguishable from normal (*23*).

The chromosomal assignment of the vWF gene was established by use of somatic hybrid mapping panels (*24*) and by in situ chromosomal hybridization (*25, 26*) (Fig. 6, A to C). The hybrid cell studies indicated that the vWF locus lies on 12p or proximal 12q (Fig. 6A). By the in situ hybridization analysis the vWF gene was localized to the region 12p12 → 12pter (Fig. 6, B and C).

Discussion. We report here the isolation of nearly full-length cDNA for human vWF and initial characterization of the vWF genetic locus. Several aspects of our cloning experiments deserve comment.

The initial vWF cDNA clone was isolated by antibody screening in the λgt11 expression vector. Only one positive clone was identified in 3×10^6 recombinant phage. This low frequency of positive plaques on initial screening of the HUVEC libraries with anti-vWF contrasts with the much higher abundance of vWF clones in the same libraries when assayed by molecular hybridization. This discrepancy is partially explained by the fact that only one of six vWF cDNA inserts is expected to be in the appropriate orientation and reading frame for expression as a fusion protein product. Additional factors are likely to contribute more substantially. Other features of the cDNA insert may influence the level of fusion protein in *E. coli*. Specifically, short peptide extensions may yield fusion products that are more stable or produced in larger amounts. In support of this hypothesis, a number of other workers using the λgt11 system have noted small inserts in their initial antibody screening isolate (*7*). In addition, epitopes expressed by cDNA clones corresponding to different regions of coding sequence may vary in reactivity with anti-vWF. Another important factor may be the characteristics of our particular anti-vWF preparation, that is, its relative avidity and titer for the multiple individual epitopes in vWF. It is possible that a different antibody preparation might detect additional clones in our cDNA library by this screening method.

Partial sequence of vWF cDNA in conjunction with primary COOH-terminal peptide sequence [our own and (*19*)], establishes that our cDNA's are derived from authentic vWF mRNA and indicates that the extensive posttranslational processing of the vWF precursor is restricted to the NH_2-terminus. DNA sequence analysis of 2.7 kb at the 5′ end of our cDNA clone pVWH33 (*23*) reveals

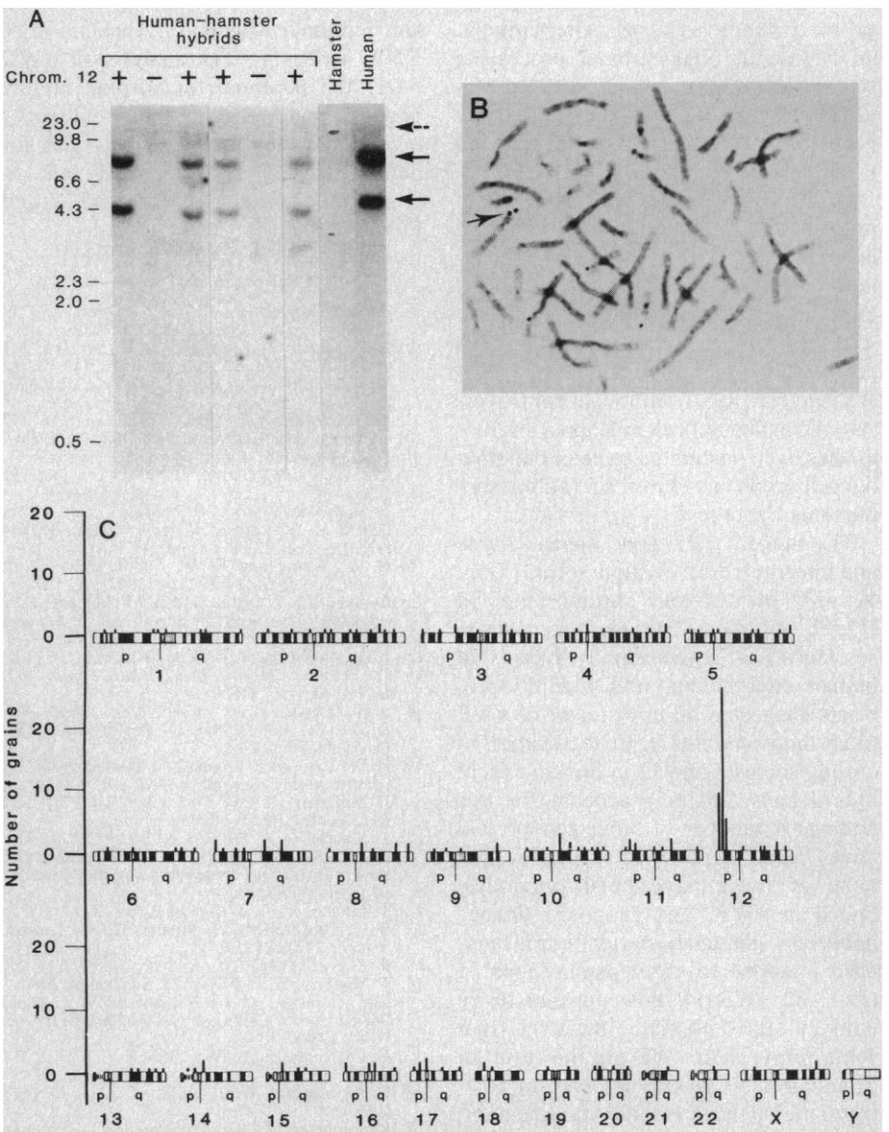


Fig. 6. Chromosomal mapping of the human vWF genetic locus. (A) Hybridization of the insert of pVWE6 with Bam HI restricted DNA's from human-hamster somatic cell hybrids (lanes 1 to 6), hamster cells (lane 7), and human cells (lane 8). The hybrid cells corresponding to the DNA's in lanes 1, 3, 4, and 6 contain human chromosome 12p or proximal 12q while those corresponding to lanes 2 and 5 lack this chromosome. The segregation of the two human genomic components recognized by this probe was completely concordant with that of chromosome 12p–proximal 12q in the 21 clones analyzed and discordant with the segregation of each of the other autosomes, the X and the Y (discordancy fractions 0.20 to 0.63). The solid arrows indicate the two human components and the dashed arrow the hamster component. Size marker positions are shown at the left. (B) In situ hybridization was performed by a modification of the method of Harper and Saunders (*25*) as previously described (*26*) with tritium-labeled pVWE6 as probe; positive grains over distal 12p are indicated by an arrow. (C) Histogram of positive grains from a total of 134 metaphases evaluated. Thirty-nine (29 percent) contained tritium over the region 12p12 → 12pter and, of a total of 192 grains, 40 (21 percent) were found over this region. No other major sites of hybridization were seen.

one continuous open reading frame. In addition, a putative hydrophobic leader sequence is present immediately after the first ATG in clone pVWH33. These data suggest that the coding region extends close to the 5′ end of the cDNA. Given the location of the termination codon at position 8.15 kb (Figs. 2 and 3), these data imply that the primary vWF translation product is at least 300 kD in size. Thus, the intracellular vWF precursor observed in pulse-chase experiments may be larger than the previous estimates of 240 to 260 kD as determined by gel electrophoresis (*1, 4*). Alternatively, an additional NH_2-terminal processing step may be required to generate the 240- to 260-kD precursor. Further studies are necessary to define these processing events.

Of the normal tissues examined, vWF mRNA was detected only in cultured endothelial cells, consistent with the cellular distribution of vWF as determined by antibody staining methods (*1, 3*). vWF mRNA of normal size was also detected in a biopsy specimen of Kaposi's sarcoma (*23*). Although controversial, the malignant cell in Kaposi's sarcoma has been postulated to be of endothelial cell origin (*3*). Our observations support this hypothesis.

The human vWF gene appears large and interrupted by multiple introns (Fig. 5) and resides on chromosome 12 (12p12 → 12pter) (Fig. 6). A report assessing vWF expression in hybrids of human endothelium and rodent fibroblasts suggested an assignment of vWF to chromosome 5 (*27*). Cosegregation of chromosomes 5 and 12 in human-rodent hybrid lines (*28*) may account for this finding. A number of other genetic loci have been mapped to 12p (*29*) but none of these have been previously linked to vWF. The proposed linkage between glutamate-pyruvate-transaminase [mapped to chromosome 8 or 16 (*29*)] and vWD (*5*) now appears to be unlikely. It is possible, however, that some subtypes of vWD are the result of genetic defects at loci involved in post-translational processing, modification, or transport of vWF. Such defects could result in a vWD phenotype that is unlinked to the vWF gene. The two patients studied here showed no evidence for a gross alteration at the vWF locus. They may have either a small deletion or insertion, or a single nucleotide substitution in the vWF gene or, alternatively, a defect at another locus as the molecular basis of their vWD.

Further studies of vWF biosynthesis and the molecular basis of vWD should benefit from the availability of vWF cDNA clones. The complete amino acid sequence of vWF, predicted from the cDNA, will help define the nature of vWF precursor processing, secretion, and multimer assembly. Application of DNA probes to the analysis of vWD offers the potential for sorting out the phenotypic heterogeneity of the disorder on a molecular genetic level, and for defining those regions of vWF that directly influence protein assembly and function.

References and Notes

1. T. S. Zimmerman, Z. M. Ruggeri, C. A. Fulcher, *Prog. Hematol.* **13**, 279 (1983).
2. J. Gitschier *et al.*, *Nature (London)* **312**, 326 (1984); W.I. Wood *et al.*, *ibid.*, p. 330; G. A. Vehar *et al. ibid.*, p. 337; J. J. Toole *et al.*, *ibid.*, p. 342.
3. L. S. Guarda *et al.*, *Am. J. Clin. Pathol.* **76**, 197 (1981); M. Miettinen *et al.*, *ibid.* **79**, 32 (1983).
4. D. D. Wagner and V. J. Marder, *J. Biol. Chem.* **258**, 2065 (1983); D. C. Lynch *et al.*, *ibid.*, p. 12757.
5. L. R. Goldin *et al.*, *Am. J. Med. Genet.* **6**, 279 (1980); M. S. Verp *et al.*, *Clin. Genet.* **24**, 434 (1983).
6. R. A. Young and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1194 (1983); in *Genetic Engineering*, J. Setlow and A. Hollaender, Eds. (Plenum, New York, in press), vol. 7; *Science* **222**, 778 (1983).
7. J. E. Schwarzbauer *et al.*, *Cell* **35**, 421 (1983); R. C. Peterson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4363 (1984).
8. S. P. Leytus *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3699 (1984); D. Foster and E. W. Davie, *ibid.*, p. 4766.
9. HUVEC's were originally obtained from M. Gimbrone [M. A. Gimbrone, Jr., R. S. Cotran, J. Folkman, *J. Cell Biol.* **60**, 673 (1974)] and were contributed by R. Weinstein.
10. T. Maciag, G. A. Hoover, R. Weinstein, *J. Biol. Chem.* **257**, 5333 (1982); bovine endothelial cell growth factor and fibronectin were both provided by R. Weinstein.
11. T. Maciag *et al.*, *J. Cell Biol.* **91**, 420 (1981).
12. S. C. Thornton, S. N. Mueller, E. M. Levine, *Science* **222**, 623 (1983).
13. R. C. Strohman *et al.*, *Cell* **10**, 265 (1977).
14. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982).
15. H. Okayama and P. Berg, *Mol. Cell. Biol.* **2**, 161 (1982).
16. U. Gubler and B. J. Hoffman, *Gene* **25**, 263 (1983).
17. Rabbit heteroantiserum prepared against human factor VIII-vWF was passed over gelatin-Sepharose to remove contaminating fibronectin, and affinity-purified by absorption and elution from a column of vWF-Sepharose. vWF-Sepharose was prepared by incubation (overnight at 40°C) of 1 mg of purified vWF with 10 ml of CNBr-activated Sepharose 4B (Pharmacia) at *p*H 7.6 and subsequent washing with 1 liter of TBS (50 m*M* tris, *p*H 8.0, 150 m*M* NaCl). The antiserum (2 ml) was applied to 1 ml of packed gel and washed extensively, and the antibody was eluted at *p*H 11. Purified vWF protein was digested with staphylococcal V8 protease, reduced, fractionated by electrophoresis on 7 percent polyacrylamide gels (PAGE) in the presence of sodium dodecyl sulfate (SDS) and β-mercaptoethanol, transferred to nitrocellulose (Western blot), and stained with the affinity-purified vWF antibody preparation (anti-vWF); an immunoperoxidase indicator system was used. Multiple bands were detected, including all those seen on a duplicate gel stained with Coomassie blue, indicating that this antibody preparation recognized multiple epitopes on the vWF molecule. This anti-vWF was used to screen the λgt11 endothelial cell cDNA libraries as described in the text. This anti-vWF was also used in a "Western blot" analysis of the β-galactosidase fusion protein product of LVWd (Fig. 1C). The blot procedure was identical to that described above except that incubation with affinity purified, ^{125}I-labeled horse antibody to rabbit Ig followed by autoradiography was used as the indicator system.
18. The single letter abbreviations for the amino acid residues are: alanine, A; arginine, R; asparagine, N; aspartic acid, D; cysteine, C; glutamic acid, E; glutamine, Q; glycine, G; histidine, H; isoleucine, I; leucine, L; lysine, K; methionine, M; phenylalanine, F; proline, P; serine, S; threonine, T; tryptophan, W; tyrosine, Y; valine, V. The single letter abbreviations for the DNA bases are A, adenine; C, cytosine; G, guanine; and T, thymine.
19. K. Titani *et al.*, paper presented at the American Heart Association National Meeting, Miami, November 1984.
20. M. E. Legaz *et al.*, *J. Biol. Chem.* **248**, 3946 (1973).
21. N. J. Proudfoot, *Nature (London)* **307**, 412 (1984).
22. J. Loscalzo, M. Fisch, R. I. Handin, *Biochemistry*, in press.
23. D. Ginsburg, D. T. Bonthron, R. I Handin, S. H. Orkin, unpublished data.
24. N. Kanda *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4069 (1983).
25. M. E. Harper and G. F. Saunders, *Chromosoma* **83**, 431 (1981).
26. T. A. Donlon *et al.*, *Am. J. Hum. Genet.* **35**, 1097 (1983).
27. N. A. Kefalides, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **40**, 327 (abstract) (1981).
28. G. P. Bruns, unpublished data.
29. P. S. Gerald and K. H. Grzeschik, *Cytogenet. Cell Genet.* **37**, 103 (1984); T. B. Shows, P. J. McAlpine, R. L. Miller, *ibid.*, p. 340.
30. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
31. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
32. J. Messing, *Methods Enzymol.* **101**, 20 (1983).
33. A. P. Feinberg and B. Vogelstein, *Anal. Biochem.* **132**, 6 (1983).

34. We thank Dr. Robert Weinstein and Karen Wenc for assistance in growing HUVEC's, Dr. Alex Markham for help with the identification of clones isolated in the library rescreening, Dr. Paula Bockenstedt for assistance with the antibody preparation and fusion protein analysis, Dr. Alan Ezekowitz for help with the Western blots, Drs. Carl O'Hara and Jerome Groopman for providing the Kaposi's sarcoma specimen, and Dr. Joseph Loscalzo and Jean-Paul Bouissel for help with the carboxypeptidase amino acid sequence analysis. Supported by NIH grants to D.G., S.H.O., R.I.H., and S.A.L., and a Junior Faculty Clinical Fellowship Award (D.G.) from the American Cancer Society.

26 April 1985; accepted 23 May 1985