Construction of cDNA coding for human von Willebrand factor using antibody probes for colony-screening and mapping of the chromosomal gene

Cornelis L.Verweij, Carlie J.M.de Vries, Ben Distel, Anton-Jan van Zonneveld, Ad Geurts van Kessel+, Jan A.van Mourik\* and Hans Pannekoek§

Department of Molecular Biology and \*Department of Blood Coagulation, Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, P.O. Box 9406, <sup>1006</sup> AK Amsterdam, and +Department of Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands

Received 15 April 1985; Revised 11 June 1985; Accepted 18 June 1985

#### ABSTRACT

Von Willebrand Factor (vWF) mRNA was identified in fractionated polyA<sup>T</sup> RNA preparations isolated from cultured human endothelial cells. Micro-injection of specific polyA<sup>'</sup> RNA fractions in Xenopus laevis oocytes provoked the synthesis of a vWF-like product which could be detected with an immunoradiometric assay relying on Sepharose-linked monoclonal anti-vWF IgG and different radiolabeled monoclonal anti-vWF IgGs. A vWF-mRNA-containing polyA<sup>T</sup> RNA preparation served as substrate for a size-selected cDNA-expression library of 60 000 colonies which was screened for the synthesis of antigens related to vWF, using polyclonal anti-vWF IgG and a second antibody conjugated with peroxidase. Eight positive colonies were detected of which two reacted strongly in the enzyme-linked assay. Immunoblotting of bacterial extracts of "expression clones" with a monoclonal anti-vWF IgG revealed polypeptides which size fits within the length of the cDNA insertions. Northern blotting of human endothelial RNA, employing fragments of vWF cDNA as probes, showed specific hybridization with a mRNA of about 9000 nucleotides. DNA-sequence analysis of a vWF-cDNA insertion revealed an open reading frame followed by a translation stopcodon. It is argued that the cDNA insertions encode the carboxy-terminal part of the vWF protein. vWF-cDNA probes were employed to map the von Willebrand factor gene on chromosome 12 using a panel of 35 human-rodent somatic cell hybrids.

#### INTRODUCTION

The glycoprotein complex Factor VIII-von Willebrand Factor (FVIII-vWF) plays a major role in hemostasis (1). It consists of two components, namely the von Willebrand Factor (vWF) and factor VIII-C (FVIII-C). The two components have distinct biological properties in hemostasis. FVIII-C is a clotting factor which has an essential function in the intrinsic coagulation pathway, whereas vWF promotes the adherence of platelets to the subendothelium after vascular injury (2-4).

FVIII-vWF circulates in plasma as a series of multimers with a relative mass  $(M_r)$  of 0.8-20x10<sup>6</sup> (1,5). The constituent unit of vWF is a glycoprotein with a  $M<sub>r</sub>$  of 225 000 which is synthesized by vascular endothelial cells as a precursor protein of about 260 000 (6-8). The precursor protein is apparently processed during secretion into the lumen to the mature vWF protein

(8,9). Assembly of the mature protein into large aggregated forms is a prerequisite for its biological function. Furthermore, vWF is considered to be the carrier protein for the procoagulant activity and apparently modulates its turnover (10).

A defective or absent vWF protein results in a bleeding disorder known as von Willebrand's disease which afflicts about 10 individuals per 100 000 people. A substitution therapy with plasma or cryoprecipitate is often adequate, but is hampered by the potential presence of infectious agents in the plasma preparations (11,12). These problems can be circumvented by using vWF protein made by recombinant DNA techniques. Such a product may also be necessary to reconstruct a stable FVIII-C preparation made by these techniques for the treatment of hemophilia-A patients (13,14). Cloning of the cDNA is the first step towards the production of recombinant vWF. A vWF cDNA will also enable us to study the structure and function of the vWF protein. Furthermore, it offers us the possibility to determine the chromosomal location of the vWF gene and analyze the structure of the vWF gene of both normal individuals and of patients with von Willebrand's disease.

### MATERIALS AND METHODS

#### DNA analysis

Plasmid DNA isolations, isolation of genomic DNA, analysis of restriction enzyme-digested DNA on agarose or polyacrylamide gels, Southern and Northern blotting and radiolabeling of DNA by "nick translation" was done as described (15). DNA sequencing was carried out by the dideoxy chain-termination procedure (16).

### Cell cultures

Endothelial cells were isolated from veins of human umbilical cords according to Jaffe et al. (17) with some modifications (18). Cells used for RNA isolation had received three or four passages and reached confluency before harvesting. After trypsinization, the cells were washed with 10 mM sodium phosphate (pH 7.4), 0.14 M NaCl and frozen in liquid nitrogen until used.

# Isolation of endothelial polyA<sup>+</sup> RNA

Total RNA was prepared from about  $2x10<sup>9</sup>$  endothelial cells, by the "hot phenol" procedure (19) with the following modifications. Lysis of the cells was done in the presence of 10 mM vanadyl ribonucleoside complex to prevent RNA degradation. Furthermore, contaminating DNA was removed by a selective precipitation of RNA in 3.3 M LiCl for 30 min at 0°C. PolyA<sup>+</sup> RNA was isolated after two cycles of adsorption to and elution from oligo-(dT) cellulose (20). The yield of polyA<sup>+</sup> RNA from  $2x10^9$  endothelial cells was approximately 300 µg. The polyA<sup>+</sup> RNA preparation was judged to be relatively undegraded, since it programmed the synthesis of discrete, long polypeptides in an in-vitro rabbit reticulocyte lysate.

# Size fractionation of polyA<sup>+</sup> RNA

 $PolyA^+$  RNA was size-fractionated by preparative urea-agarose gel electrophoresis according to the procedure of Rosen et al. (21) except that low melting agarose (1.5%) was used. After electrophoresis, the gel was cut into slices and RNA fractions were extracted as described (22). Alternatively, polyA<sup>+</sup> RNA was fractionated according to its size by centrifugation, employing linear denaturing sucrose gradients. Sucrose gradients (5-30%) were made in 0.1 M Tris-HCl (pH 7.5), <sup>1</sup> mM EDTA, 0.2% SDS and 50% (v/v) formamide. RNA samples in the same buffer were heated for 2 min at 55°C and cooled in ice before layering onto the gradients. Centrifugation was performed in a SW41 rotor at 15°C for 15 h at 85 OOOxg.

## Microinjection of RNA into Xenopus laevis oocytes

Microinjection of RNA into Xenopus laevis oocytes was performed essentially according to Gurdon et al. (23). Twenty five nl were injected of either total endothelial polyA<sup>+</sup> RNA, size-fractionated endothelial polyA<sup>+</sup> RNA, tobacco-mosaic virus (TMV) RNA or calf  $\alpha$ -crystalline RNA. All RNA preparations were adjusted to <sup>1</sup> mg/ml before injection. Microinjected oocytes were incubated for 30 h at 19°C in modified Barth's medium (23). Subsequently, pools of 10 undamaged oocytes were homogenized in 0.25 ml 10 mM Tris-HCl (pH 9.0), 50 mM NaCl and centrifuged for 30 min in a microfuge. Lipids were removed from the supernatant by two successive extractions with an equal volume of diethylether. The presence of vWF-like polypeptides was detected with an immunoradiometric assay.

Immunoradiometric assay (IRMA) for the detection of vWF-like products An immunoradiometric assay (IRMA) was used relying on Sepharose-linked monoclonal anti-human vWF CLB-RAg 20 and a mixture of six different <sup>125</sup>Ilabeled anti-human vWF IgGs (CLB-RAg 20, CLB-RAg 23, CLB-RAg 35, CLB-RAg 38, CLB-RAg 56, CLB-RAg 58), representing 2x104 counts per min. The incubation conditions and washing procedures were as described (24). The lower limit of detection of this assay is about 5 ng (0.25 mU) of vWF-like protein. The procedure probably allows a quantitative recovery of vWF-like products, because administration of <sup>a</sup> standard amount of purified vWF to homogenized uninjected oocytes showed full recovery of vWF-related antigen.

### cDNA synthesis and bacterial transformation

 $cDNA$  synthesis, using unfractionated endothelial polyA<sup>+</sup> RNA as substrate. was performed essentially according to Gubler and Hoffman (25). cDNA was size-fractionated employing Sepharose CL-4B chromatography and cDNA smaller than 600 bp was discarded. Typically, 600 ng of polyA<sup>+</sup> RNA vielded approximately 200 ng size-fractionated, double-stranded "C-tailed" cDNA which was annealed to a twofold molar excess of G-tailed PstI-cleaved pUC9 (26). Transformation of strain E.coli DH1 (recA) was carried out by the high efficiency procedure of Hanahan (27).

## Screening on expression of vWF-like products in E.coli

Duplicate filters, each carrying approximately 3000 colonies, were exposed for 20 min to an atmosphere saturated with vapor of chloroform (28). Then, each filter was incubated separately for 16 h at room temperature in 20 ml 50 mM Tris-HCl (pH 7.5), 0.15 M NaCl, 5 mM MgCl<sub>2</sub>, 0.1% Tween-20, 1  $\mu$ g/ml pancreatic DNaseI and 40 pg/ml lysozyme under gentle agitation. Subsequently, the filters were rinsed in 50 mM Tris-HCl (pH 7.5), 0.15 M NaCl (TS), carefully polished with tissue paper and rinsed again in TS. The filters were then incubated for <sup>1</sup> h at room temperature in TS containing 0.5% Tween-20, followed by a 10-min rinse in TS with 0.1% Tween-20. The incubation of fixed bacterial antigens with the first antibody preparation was for <sup>1</sup> h at room temperature, using a 1:500 dilution in TS, 0.1% Tween-20 of rabbit-anti-human vWF IgG which had been preabsorbed for 2 h at 4°C with a boiled extract of E.coli DH1 bacteria. After the antigen-antibody reaction, the filters were washed five times in TS, 0.1% Tween-20 and subsequently incubated for <sup>1</sup> h at room temperature with a 1:1000 dilution of horse-antirabbit IgG conjugated to horse-radish peroxidase. Finally, the filters were washed five times with TS, 0.1% Tween-20, followed by one wash step in TS, 0.5% Tween-20 and stained for peroxidase activity.

#### In-situ immunostaining of plasma proteins

SDS-agarose electrophoresis and fixation was performed as described (29) except that a 2.5% agarose gel was used. Subsequently, the gel was incubated for at least <sup>1</sup> h at room temperature in TS, 3% bovine serum albumin, 0.1% gelatin, 0.6% human IgG, 0.05% Tween-20, 0.02% (w/v) sodium azide and then incubated for 16 h at room temperature with 1  $\mu$ g/ml of the murine anti-human vWF IgG CLB-RAg 41. This incubation was followed by extensive washing of the gel for <sup>1</sup> h in 0.5 M NaCl and for 8 h in TS, respectively. Antigen-antibody complexes are detected after incubation for 16 h at room temperature with <sup>125</sup>I-labeled goat-anti-mouse IgG, which was affinitypurified on a column containing covalently bound murine IgG. Subsequently, the gel was washed extensively as described above. Finally, the gel was dried and autoradiography was performed.

## Immunoblotting of bacterial proteins

The analysis by immunoblotting of bacterial proteins was done after SDSpolyacrylamide gel electrophoresis according Lo Laemmli (30) and transfer of proteins from the gel into nitrocellulose membrane filters essentially as done before (31). Bacterial extracts were prepared as described (32). Finally, the proteins were solubilized in 0.5 ml loading buffer (30), heated for 5 min at 100°C and directly loaded onto the gel (25  $\mu$ 1 per slot). Polypeptides blotted on nitrocellulose filters were incubated with a 1:400 dilution of affinity-purified rabbit polyclonal anti-human vWF IgG or separately with 21 different murine monoclonal anti-human vWF IgG's  $(1 \mu g)$ ml). Immunocomplexes were detected either with <sup>125</sup>I-labeled sheep-antirabbit IgG or  $125I$ -labeled affinity-purified goat-anti-mouse IgG  $(2x10^6$ counts per min). Incubation, blocking of the filters and washing was the same procedure as described for colony-screening in a previous paragraph, except that for these blots 0.2% gelatin was added to all buffers and for the screening with monoclonals, the blots were finally washed for 16 h with 50 mM Tris-HCl (pH 7.5), 0.5% Tween-20, 0.5 M NaCl, 0.2% gelatin. After washing, autoradiography was performed.

### Materials

Immunoreagents were either from the Central Laboratory of the Netherlands Red Cross Blood Transfusion Service or from Dakopatts (Copenhagen, Denmark). "Super" reverse transcriptase was obtained from Anglia Biotechnology (UK), whereas restriction endonucleases were from New England Biolabs and used according to the manufacturers' instructions. Monoclonal antibodies were purified from ascites by absorption to a protein-A Sepharose column (Pharmacia Fine Chemicals, Uppsala, Sweden). <sup>125</sup>I-labeling of polyclonal IgG was carried out by the chloramine-T method, whereas monoclonal IgG preparations were radiolabeled using the Iodogen reaction.

### RESULTS

### Identification of vWF-mRNA in endothelial cells

We intend to study the structure and function of human endothelial proteins involved in hemostasis and thrombosis. For that purpose, we have constructed cDNA clones, programmed by endothelial polyA<sup>+</sup> RNA. In this paper, we focus on the construction and isolation of vWF-cDNA clones.

The substrate for these cDNA clones, vWF-mRNA, was identified in endo-



Figure 1. Identification of  $vWF$  mRNA. Endothelial polyA<sup>+</sup> RNA (150 µg) was subjected to electrophoresis on a preparative 1.5% urea-agarose gel. The gel was sliced yielding 25 fractionş and RNA was isolated as described in<br>Materials and Methods. Total polyA<sup>'</sup> RNA (pA) and RNA fractions migrating slower than the 18S ribosomal RNA marker were injected into Xenopus laevis oocytes. As controls, oocytes were injected with calf  $\alpha$ -crystalline RNA (a-crys) or with tobacco-mosaic-virus (TMV) RNA. Homogenates of injected oocytes were tested for de novo synthesis of vWF-like products using an immunoradiometric assay (IRMA) as outlined in Materials and Methods. The arrows in the right panel indicate the positions of ribosomal RNA's run in a parallel lane.

thelial cells. PolyA<sup>+</sup> RNA preparations were shown to contain intact vWFmRNA by demonstrating specific translation of microinjected RNA in Xenopus laevis oocytes. These translation products, i.e. vWF-like polypeptides, were detected using an immunoradiometric assay (IRMA) (24). Extracts of oocytes were incubated with monoclonal antibodies (anti-vWF IgG CLB-RAg 20) coupled to Sepharose beads and subsequently with a mixture of six different monoclonal antibodies (anti-vWF IgGs) which had been labeled with 125I. The amount of Sepharose-bound 125I is indicative of the amount of vWF-like products synthesized by microinjected oocytes. An analysis was performed using both total polyA<sup>+</sup> RNA and polyA<sup>+</sup> RNA size-fractionated by urea-agarose gel electrophoresis, a technique with a high resolution for heterogeneous RNA preparations. The data are given in Fig. 1.

A significant amount of Sepharose-bound 125I-labeled material was detected using extracts of oocytes injected either with total polyA<sup>+</sup> RNA from endothelial cells or with mRNA fractions with a length considerably larger than the marker 28S ribosomal subunit. These results indicate that our

+ Nucleic Acids Research<br>
The Company of the polyA RNA preparations contain mRNA which encode vWF-like polypeptides. We repeated the experiment described above with polyA<sup>+</sup> RNA size-fractionated by sucrose/formamide centrifugation which technique allows a more reliable RNA length determination than urea-agarose electrophoresis (33). Essentially, the same results were obtained as for  $polyA^{\dagger}$  RNA size-fractionated by urea-agarose gel electrophoresis. Only products derived from polyA<sup>+</sup> RNA with a length of 7500 to 10 000 nucleotides bound significantly more  $125I$ labeled monoclonal anti-vWF IgG than control mRNA.

This result is in accord with the observation that the precursor protein for the vWF subunit has a molecular weight of approximately 260 000 daltons (8). Fifteen per cent of the molecular weight of this glycoprotein is accounted for by carbohydrate residues (34). Hence, the coding region required should comprise at least 6000 nucleotides.

## Construction and isolation of vWF-cDNA clones

The availability of well-characterized polyclonal and monoclonal anti-vWF antibodies (35) allowed us to use a technique for screening bacterial colonies based on the expression of antigenic determinants (28). For that purpose, we constructed a human endothelial cDNA library, using E.coli as a host and plasmid pUC9 as vector (26). pUC9 contains the lac promotor, the translational initiation signals and 5 N-terminal aminoacids of  $\beta$ -galactosidase, followed by multiple unique cloning sites (e.g. PstI). An expression library of 60 000 independent colonies was constructed of which approximately 90% contained plasmids with cDNA insertions with a length longer than 600 bp. This library was screened for bacterial synthesis of vWF-like proteins. After lysis of the colonies, the bacterial proteins were incubated with a rabbit polyclonal anti-human vWF IgG preparation and subsequently with sheep-anti-rabbit IgG conjugated with horse-radish peroxidase. This enzyme-linked immunosorbent assay (ELISA) revealed eight colonies which were positive on duplicate filters. An example is given in Fig. 2. Rescreening consistently yielded positive signals. Moreover, rescreening with an affinity-purified rabbit polyclonal anti-human vWF IgG prepared from a different serum gave identical results. It should be noted that two colonies reacted strongly in the ELISA, two colonies gave an intermediate signal, whereas the remaining four colonies reacted weaker in the ELISA.

An analysis of plasmid DNA from these eight immunoreactive clones showed that the length of the DNA insertions ranged from 750 to 1415 bp. Clones which reacted weakly in the screening appeared to be similar, based on re-



Figure 2. Immunological detection of a vWF-cDNA clone. Colonies on duplicate filters were screened with rabbit-anti-human vWF IgG, followed by incubation with horse-anti-rabbit IgG conjugated with horse-radish peroxidase after which reactions the filters were stained for peroxidase activity. A detailed procedure is given in Materials and Methods. The positive colony on the duplicate filters (indicated by the arrow) contained a plasmid with a 1210 bp cDNA insertion. The crosses on the filters are marking the orientation.

striction endonuclease analysis. The cDNA insertions of these clones all spanned approximately 1415 bp. In contrast, the remaining clones contained plasmid DNA with insertions which were different in length, namely 750, 780, 830 and 1210 bp. These plasmids will be referred to as pvWF 750, pvWF780, etc. By Southern blot analysis of restricted pvWF-DNA preparations, using restriction fragments of different insertions as probes, it was shown that all vWF-cDNA insertions consist of overlapping nucleotide sequences.

The vWF-like proteins, synthesized by the two strong expression clones pvWF1210 and pvWF780, were characterized by immunoblotting using an affinity-purified rabbit polyclonal anti-human vWF IgG preparation. The results, given in Fig. 3A, show that these two plasmids encode different discrete  $vWF$ -like proteins. Plasmid pvWF1210 codes for a product with a M<sub>r</sub> of 37 000, whereas plasmid pvWF780 codes for a product with a  $M_r$  of 20 000. A schematic representation of the length and composition of these two vWFcDNA inserts and the encoded proteins is presented in Fig. 4.

The formation of specific immunocomplexes between a bacterial antigen and (a) monoclonal anti-human vWF IgG would provide additional evidence for



Figure 3. Imnunoblotting of protein extract of bacteria containing pvWF-DNA with polyclonal and monoclonal anti-vWF IgG (A and B) and in situ immunostaining of plasma proteins with a monoclonal anti-vWF IgG (C). (A) Protein extracts transferred to nitrocellulose after electrophoresis on a 13% SDSpolyacrylamide gel were incubated with rabbit polyclonal anti-vWF IgG (affinity purified on a Factor VIII-von Willebrand factor column) followed by an incubation with 125I-labeled sheep-anti-rabbit IgG. Extracts were prepared from strain E.coli DHI containing: lane 1, vector pUC9 without insert; lane 2, pvWF780, and lane 3, pvWFI210. (B) Protein extracts were subjected to electrophoresis on preparative 13% SDS-polyacrylamide gels and transferred to nitrocellulose sheets. Sections of the blots, perpendicular on the "protein bands", were incubated with affinity-purified rabbit polyclonal anti-vWF IgG and <sup>a</sup> panel of different monoclonal anti-vWF IgGs. Subsequently, these blots were incubated with  $125$ I-labeled sheep-anti-rabbit IgG and affinity-purified goat-anti-mouse IgG, respectively. The essential parts of these blots are shown in this figure. Lanes 1-5 contain protein extracts of E.coli DHI harboring pvWF78O and lanes 6-10 E.coli DH1 with pvWF1210. The proteins in lanes <sup>1</sup> and 6 are incubated with affinity-purified rabbit polyclonal anti-vWF IgG as a reference. Proteins in lanes 2 and <sup>7</sup> are incubated with monoclonal anti-vWF CLB-RAg 42, in lanes 3 and 8 with CLB-RAg 41, in lanes 4 and 9 with CLB-RAg 35 and in lanes 5 and 10 with CLB-RAg 23. (Exposure time for the lanes 2-5 and 7-10 was 15 times longer than for the lanes 1 and 6.) (C) In situ immunostaining of Factor VIII-von Willebrand factor. Plasma samples were pretreated at  $37^{\circ}$ C with SDS at a final concentration of 2% and subjected to electrophoresis on a 2.5% agarose gel containing 0.1% SDS. After fixation, the gel was incubated with murine monoclonal antihuman vWF CLB-RAg 41 IgG. Immunocomplexes were detected with 12SI-labeled goat-anti-mouse IgG. Lane 1, 10 µl plasma of a normal individual. Lane 2, 10 p1 plasma of a patient with severe homozygous von Willebrand's disease. Incubation of proteins from normal plasma without monoclonal anti-human vWF CLB-RAg 41 did not show staining with 125I-labeled goat-anti-mouse IgG.



Figure 4. Schematic representation of the structure of the vWF-cDNA inserts of the two strong expression clones pvWF780 and pvWFI210. The top line represents vWF-mRNA. The translation termination codon is indicated by <sup>a</sup> black dot. The lines below vWF-mRNA represent the vWF-cDNA inserts of the two clones. The expressed vWF-like proteins are indicated by wavy lines below the corresponding coding part of vWF cDNA.

the authenticity of plasmids harboring vWF-cDNA. For that purpose, we have decided to further investigate the vWF-like proteins encoded by these plasmids. Bacterial extracts of E.coli DHI containing either pvWF1210 or pvWF780 were run on SDS-polyacrylamide gels, blotted onto nitrocellulose paper and incubated separately with several different murine monoclonal anti-human vWF IgGs. Subsequently, immunocomplexes were detected by incubation with radiolabeled goat-anti-mouse IgG. The results are shown in Fig. 3B. It is evident that our panel of monoclonals harbors one antibody preparation (CLB-RAg 41) which forms <sup>a</sup> specific immunocomplex with the bacterial vWF-like proteins encoded by these two plasmids. At this point, it should be emphasized that the vWF-like products, encoded by pvWFl210 and pvWF780, react identically with polyclonal and monoclonal anti-vWF sera in immunoblotting experiments (3B). In this respect, it is important to define the properties of monoclonal CLB-RAg 41. Previously, it has been shown with IRMAs that monoclonal CLB-RAg 41 binds to purified FVIII-vWF, to FVIII-vWF present in normal plasma, but not to a component in vWF-deficient plasma and, finally, caused an inhibition of the adherence of radiolabeled platelets to the subendothelium (35). Furthermore, monoclonal anti-vWF antibodies display a typical granular immunofluorescence pattern with endothelial cells, revealing discrete organelles presumably identical with so-called Weibel-Palade bodies (36). Treatment of the cells with specific



agents (e.g. thrombin) causes release of vWF and disappearance of fluorescence. Identical results are obtained with monoclonal anti-vWF CLB-41 (not shown), demonstrating that CLB-41 reacts only with vWF and not with other endothelial proteins. In Fig. 3C, we show that monoclonal anti-human vWF CLB-RAg 41 specifically binds to SDS-denatured FVIII-vWF from normal plasma and not to a component present in plasma from a patient with severe homozygous von Willebrand's disease. Multimeric forms of FVIII-vWF present in normal plasma show a characteristic pattern (1,29). These observations demonstrate that monoclonal anti-vWF CLB-RAg 41 specifically binds to both human FVIII-vWF and to bacterial proteins encoded by plasmids pvWF1210 and pvWF780, as shown in Fig. 3. Hence, these polypeptides from different origin harbor the same epitope. From the results presented so far, we conclude that we have cloned part of the human vWF gene, which codes for a bacterial protein, constituting part of the native human vWF protein.

To determine the length of vWF mRNA, both total and polyA<sup>+</sup> RNA preparations from endothelial cells were subjected to electrophoresis on denaturating agarose gels and blotted on nitrocellulose filters. Subsequently, hybridization was carried out with the nick-translated 770 bp PstI-cDNA fragment of plasmid pvWF1210. The results, presented in Fig. 5, show that

E<br>International and the serve of the ser CTG CAG TAT GTC AAG GTG GGA AGC TGT AAG TCT GAA GTA GAG GTG GAT ATC CAC TAC TGC leu gin tyr val lys val gly ser cys lys ser glu val glu val asp ile his tyr cys 120 CAG GGC AAA TGT GCC AGC AAA GCC ATG TAC TCC ATT GAC ATC AAC GAT GTG CAG GAC CAG gin gly lys cys ala ser lys ala met tyr ser ile asp ile asn asp val gin asp gin II lys val gly ser cys lys ser glu val glu val asp ile his tyr cys<br>IT GCC AGC AAA GCC ATG TAC TCC ATT GAC ATC AAC GAT GTG CAG GAC CAG<br>I20<br>ISO sala ser lys ala met tyr ser ile asp ile asn asp val gln asp gln<br>ISO sala ser ly TGC TCC TGC TGC TCT CCG ACA CGG ACG GAG CCC ATG CAG GTG GCC CTG CAC TGC ACC AAT cys ser cys cys ser pro thr arg thr glu pro met gln val ala leu his cys thr <u>asn</u> I GGA AGC TGT AAG TCT GAA GTA GAG GTG GAT ATC CAC T, gly ser cys lys ser glu val glu val asp ile his the star can get tyr ser ile asp ile as a sp val glu and a sp ile his the star can get tyr ser ile asp ile as a sp val gln 240 GGC TCT GTT GTG TAC CAT GAG GTT CTC AAT GCC ATG GAG TGC AAA TGC TCC CCC AGG AAG gly ser val val tyr his glu val leu asn ala met glu cys lys cys ser pro arg lys

TGC AGC AAG TGA GGCTGCTGCAG cys ser lys

Figure 6. Nucleotide sequence of the 260 bp PstI-cDNA fragment of pvWF1210. The 260 bp PstI vWF-cDNA fragment was subcloned in both orientations in phage M13mp9 RF DNA. The complete nucleotide sequence in both orientations was determined according to the procedure of Sanger et al. (16). In this figure, the nucleotide sequence of the coding strand is depictured and the predicted aminoacids are indicated. The presented nucleotide sequence apparently represents the C-terminal coding-sequence of the vWF gene, including the genuine TGA-translation termination codon. The nonsense codons in the other two reading frames are indicated with a line above those codons. The underlined aminoacid sequence contains a potential asparaginelinked glycosylation site.

a single mRNA species of about 9000 nucleotides hybridizes to the vWF-cDNA probe. The length of this mRNA is in accord with the estimated coding segment of approximately 6000 nucleotides and is in agreement with the results of the vWF-mRNA identification experiment shown in Fig. 1.

Part of the DNA sequence of the cDNA insertion of plasmid pvWF1210 has been determined, employing the dideoxytriphosphate chain termination method (16). Hereto, the 260 bp PstI-cDNA fragment of pvWF1210 has been subcloned into M13mp9 DNA in both orientations and the DNA sequence of both strands has been clarified. The data are given in Fig. 6. The 260 bp PstI-cDNA fragment is flanked by PstI-cDNA fragments of 770 and 180 bp, respectively. This conclusion has been drawn from the finding that the 770 and the 180 bp fragments harbor a 10-20 G-C stretch, created by terminal transferase, at one terminus, in contrast to the 260 bp fragment (data not shown). Restriction-enzyme mapping was used to establish i) that the 770 bp fragment is located upstream of the adjacent 260 bp fragment, whereas the 180 bp fragment is downstream, and ii) the orientation of the 260 bp fragment relative to the other PstI vWF-cDNA fragments. Inspection of possible reading frames revealed the following characteristics. Two reading frames

in somatic cell hybrids																									
	Chromosome Complement																								
Hybrid	vWF	1	$\overline{\mathbf{c}}$	3	4	5	6	$\overline{ }$	8	9	10	11			12 13 14 15		16 17		18	19	20	21	22	x	Y
A3RS- 5B	۰			٠	R						٠	R						٠							
A3RS-11B	٠				R						t														
A3RS-13A	٠																								
A3RS-17A	÷			÷																					
A3RS-17B				٠														٠	ŧ	ŧ	ŧ				
A3RS-27B	٠			٠							٠	R					٠	$\ddot{}$	$\ddotmark$	$\ddotmark$	÷				
A3RS-30A					R													٠							
A3Bi- 7C																		٠			٠				۰
A3Bi- 9B																									
12CB- 4B			t								t														
12CB-7B										R															
12CB-23B																									
12CB-26A																									
12CB-30A	٠																	٠							
12CB-33B		٠							٠			۰					÷					٠	$\ddot{}$		
$4iRO-6$	۰						R			٠	t			t	t						٠		۰	۰	
X3/F3																									
X3/010				R																			٠		
A19A1																									
A3MS-21F	٠																						٠		
$54 - 7 - 1$																								٠	
33mp																									
33.2																							٠	٠	
33.9																									
81.10.1		ŧ																							
81.10.7																									
81.10.18																									
81.13.1	٠																								
81.13.1BII	÷																								
81.13.3H	٠	٠																							
81.13.4H																									
81.13.5																									
WEGL <sub>i</sub> - $3^*$									ŧ	R		R											٠		
$WECLi-9$			±						ŧ														R		
$WEGLi-11$ <sup>*</sup>	۰							ŧ															R		
% Discordancy		43	51	41	44	51	44	43	40	39	40	55	0	31	34	40	46	49	57	31	46	46	52	43 54	

TABLE <sup>1</sup> Segregation of cloned vWF-cDNA with human chromosomes

Symbols: +, present; -, not present; ±, present in less than 10% of the cells tested; \*, human-mouse hybrid cells; R, rearranged. Rearranged chromosomes were not included in the scoring.

Concordant hybrids are those that have either retained or lost vWF and a specific human chromosome. Discordant hybrids are those that either retained the gene but not a specific chromosome, or the reverse. % Discordancy indicates the degree of discordant segregation for a marker and a chromosome, Lack of discordancy demonstrates chromosome assignment.



Figure 7. Southern hybridization of the 770 bp (A) and 260 bp (B) PstI vWF-cDNA probes derived from pvWFl2lO to TaqI-restricted DNA from: lane 1, Chinese hamster cell-line A3; lane 2, hybrid X3/DIO; lane 3, hybrid A19AI; lane 4, hybrid A3MS-21F; lane 5, hybrid l2CB-4B; lane 6, hybrid 12CB-23B; lane 7, hybrid l2CB-26A; lane 8, Chinese hamster cell-line E36; lane 9, hybrid 54-7-1; lane 10, hybrid 33mp, and lane 11, HeLa cells. The blots were washed under stringent conditions [65°C; 15 mM sodium chloride, 1.5 mM sodium citrate (0.1 <sup>x</sup> SSC)]. A selection of the digestions of different hybrid DNAs was done and used to compose this figure.

contain several nonsense codons which are mainly located near the <sup>5</sup>' end of the 260 bp vWF-cDNA fragment. The remaining reading frame harbors one stopcodon (TGA; 250-252) which is positioned near the <sup>3</sup>' end of this fragment. The data on the size of the bacterial vWF-like protein, encoded by plasmid pvWFl2lO (Fig. 4), strongly indicate that this reading frame is the correct one for the vWF protein. The expressed vWF-like protein of plasmid pvWFl2lO has <sup>a</sup> molecular weight of <sup>37</sup> <sup>000</sup> daltons which size corresponds with <sup>a</sup> coding region of about <sup>1000</sup> nucleotides. Obviously, the upstream located <sup>770</sup> bp PstI vWF-cDNA fragment does not suffice to encode this protein. Almost the entire coding capacity of the adjacent 260 bp PstI vWF-cDNA fragment should be translated to synthesize <sup>a</sup> protein with <sup>a</sup> molecular weight of <sup>37</sup> 000 daltons. This interpretation positions the termination codon of the vWF gene indeed in the close proximity of the <sup>3</sup>' end of the 260 bp PstI vWF-cDNA fragment. We conclude that we have cloned cDNA which encodes the carboxy-terminal part of the human vWF protein.

### Chromosome localization of the vWF gene

To map the vWF gene on a human chromosome, we used 32 human-Chinese-hamster and 3 human-mouse hybrid cell lines. All these cell lines contain a full complement of either Chinese-hamster or mouse chromosomes and a limited number of human chromosomes. Determination of the human chromosome constitution (Table 1) was performed by karyotyping of at least 16 cells, in some cases in combination with an analysis of isozymes characteristic of a human chromosome. DNA from these cells digested with TaqI was analyzed on Southern blots hybridized with the 770 bp PstI vWF-cDNA fragment of pvWF1210. A representative experiment is shown in Fig. 7. As controls, hybridization of this probe with Chinese-hamster DNA (lanes <sup>1</sup> and 8) and Hela DNA (lane 11) is shown.

Two different Chinese-hamster cell lines, designated A3 and E36, have been used for the construction of the hybrids. These two cell lines exhibit a TaqI polymorphism with this probe which is illustrated by three crosshybridizing hamster fragments (lanes 1-7) and two hamster fragments (lanes 8-10), respectively. In addition, four human-specific TaqI fragments of 1.5, 2.2, 3.1 and 3.6 kb are seen in HeLa DNA (lane 11). The 3.6-kb fragment has the same size as one of the detected Chinese-hamster fragments. Consequently, this fragment could not be identified in DNA from hybrids derived from the Chinese-hamster cell line containing this fragment. Taken this into account, the human TaqI fragments are seen in lanes 2, 4, 5 and 10. None of the human-specific TaqI fragments mentioned are present in the lanes 3, 6, 7, 8 and 9. Similar results were found with DNA digested with other restriction endonucleases. As shown in Fig. 7B, the same results were also obtained after hybridization of the blot shown in Fig. 7A with the 260 bp PstI vWF-cDNA fragment of pvWF1210. In this way, a single fragment of approximately 4.2 kb, present in human DNA, hybridized to the probe, whereas no fragment in Chinese-hamster DNA was detected.

All the results from the Southern blot-analysis studies are summarized in Table <sup>1</sup> and show that the vWF cDNA hybridizes exclusively to DNA from human chromosome 12.

#### DISCUSSION

In this paper, we report on the construction and isolation of cDNA clones which encode part of the human vWF subunit of the factor VIII-von Willebrand factor complex. The availability of well-characterized polyclonal and monoclonal anti-vWF antibodies allowed us to screen on expression of

antigenic determinants in E.coli cells. Several lines of evidence have been used to show that these overlapping cDNA insertions correspond to coding regions for the vWF protein, namely:

- ( i ) The expressed antigens are recognized by two different polyclonal anti-vWF sera, one of which had been affinity-purified on a column containing covalently bound factor VIII-von Willebrand factor.
- ( ii) Well-characterized monoclonal anti-vWF IgG (CLB-RAg 41) reacted with the expressed vWF-like polypeptides of a number of the immunoreactive clones, indicating that antigenic determinant(s) present on these bacterial products and on native vWF are identical.
- (iii) The cDNA-containing plasmids specifically hybridize with an endothelial mRNA with a length of about 9000 nucleotides. The length of this mRNA is sufficient to encode this glycoprotein with a molecular weight of 260 000 daltons. mRNA with a corresponding size also provokes the synthesis of vWF-like polypeptides in an in-vivo translation assay (Fig. 1).

From these results, we conclude that we have cloned part of the human vWF cDNA. This conclusion has been confirmed by the data of Titani and coworkers who have determined the aminoacid sequence of the carboxy-terminal cyanogen-bromide-cleaved fragment of native vWF (personal communication). Their sequence of 11 aminoacids is identical with our predicted carboxyterminal aminoacid sequence, shown in Fig. 6. This also indicates that the vWF precursor protein is not extended at its carboxy-terminus.

We have shown that the use of monoclonal antibodies can provide a rapid and convenient way to confirm the presence of, or to identify a foreign gene segment expressed in bacteria. However, it should be noted that both polyclonal and monoclonal antibody preparations used in such experiments must have sufficient affinity to reveal 0.1-1 ng of antigen synthesized in a bacterial colony and should display a low background with E.coli lysate. A comparison of the signal produced by the two strong expression clones, containing pvWF780 or pvWF1210 with a standard series of spotted native vWF, showed that the colony signal corresponds to about 2 ng of vWF-like protein. From these data, it can be calculated that approximately  $6x10<sup>4</sup>$ vWF-like molecules are synthesized per bacterium, whereas the clone harboring pvWF1210 produces about twofold less. We have not explored the reason for the different levels of expression of the vWF clones. In particular, four clones could be detected only by colony-screening, but not by immunoblotting. Possibly in these cases, the vWF-like bacterial product is unstable.

We have determined, both by electrophoresis in formaldehyde-agarose gels and after denaturation of the RNA with glyoxal and dimethyl sulfoxide, that vWF-mRNA spans about 9000 nucleotides. The coding region required for a glycoprotein of 260 000 daltons, containing 15% carbohydrate residues, will be approximately 6000 bp. Preliminary sequence analysis indicates that vWF-mRNA contains a relatively short 3'-untranslated region of about 130 nucleotides, followed by a polyA-tail of 85 residues. Hence, it is conceivable that the actual vWF precursor protein is considerably larger than 260 000 daltons, provided that vWF-mRNA does not harbor an extensive 5'-untranslated region. Presently, we are attempting to clarify this point.

Screening the cDNA library, under stringent hybridization and washing conditions, with the fully codogenic 770 bp PstI cDNA fragment of pvWF1210 revealed 40 hybridizing colonies, including the eight immunoreactive ones. it is conceivable that those 32 novel clones contain vWF-cDNA insertions which are either in the opposite orientation versus the lac promotor or are not positioned in the reading frame of the N-terminal aminoacids of  $\beta$ -galactosidase. The longest "non-expression" clone so far examined appeared to harbor a plasmid with a cDNA insertion of 2280 bp. Southern blot analysis confirmed that the insertions consist of vWF cDNA. Thus, pvWF plasmid-containing clones represent only 0.07% of the entire endothelial cDNA library, in spite of the fact that cDNA smaller than 600 bp has been discarded.

Using labeled vWF-cDNA probes and a series of human-hamster and humanmouse somatic-cell hybrids containing reduced numbers of human chromosomes, we have been able to assign the human vWF gene exclusively to chromosome 12. Hybridization of restricted DNA of different hybrids with the 770 bp vWFcDNA probe resulted in the same set of human fragments for each restriction endonuclease used. Only one TaqI-restriction fragment is detected with the 260 bp vWF-cDNA probe. Hence, it is likely that there is only one copy of the vWF gene present in the human genome. Based upon an immunological approach, Kefalides (37) concluded that the vWF gene may be located on human chromosome 5. To correlate his results with those of Edgell et al. (38) who reported suppression of vWF expression in endothelial cell hybrids, he proposed that a particular combination of either human chromosomes or human plus rodent chromosomes is required for vWF to be expressed. That hypothetical combination might have been different in the hybrids used in those two studies. Hybridization of digested DNA from a panel of somatic-cell hybrids with a well-defined cDNA probe is apparently a more reliable tech-

nique to map a gene than innunofluorescent methods. Our data clearly demonstrate that the vWF gene is not located on chromosome 5, but exclusively on chromosome 12.

Currently, we are constructing pvWF plasmids which extend into the 5' direction of vWF-mRNA. Assembly of a full-length vWF cDNA will enable us to study structure and function of the vWF protein.

#### ACKNOWLEDGEMENTS

We are grateful to Mrs. M.C. Janssen, Mr. M. Gonsalves and Dr. Ch. Willems for culturing the endothelial cells we used. We thank Dr. M.H. Salden for making us familiar with the oocyte microinjection technique, Dr. R. Verhoeff-Fremery for the Xenopus laevis toads she provided, R.F. Evers for his excellent technical assistance, Dr. P. Borst for constructive criticism during this work and Mrs. M.C.F. Theuns-Jongen for typing the script. This study was supported by the Netherlands Organization for the Advancement of Pure Research (ZWO) (grant no. -13/90-91).

§To whom correspondence should be addressed

#### REFERENCES

- 1. Hoyer, L.W. (1981) Blood 58, 1-13.
- 2. Tschopp, T.B., Weiss, H.J., and Baumgartner, H.R. (1974) J. Lab. Clin. Med. 83, 296-300.
- 3. Sakariassen, K.S., Bolhuis, P.A., and Sixma, J.J. (1979) Nature 279, 636-638.
- 4. Bolhuis, P.A., Sakariassen, K.S., Sander, H.J., Bouma, B.N., and Sixma, J.J. (1981) J. Lab. Clin. Med. 97, 568-576.
- 5. Mikaelsson, M.E., Forsman, N., and Oswaldsson, U.M. (1983) Blood 62, 1006-1015.
- 6. Jaffe, E.A., Hoyer, L.W., and Nachman, R.L. (1973) J. Clin. Invest. 52, 2757-2764.
- 7. Nachman, R.L., Levine, R., and Jaffe, E.A. (1977) J. Clin. Invest. 60, 914-921.
- 8. Wagner, D.D., and Marder, V.J. (1983) J. Biol. Chem. 258, 2065-2067.
- 9. Lynch, D.C., Williams, R., Zimmerman, T.S., Kirby, E.P., and Livingston, D.M. (1983) Proc. Natl. Acad. Sci. USA 80, 2738-2742.
- 10. Tuddenham, E.G.D., Lane, R.S., Rotblat, F., Johnson, A.J., Snape, T.J., Middleton, S., and Kernoff, P.B.A. (1982) Brit. J. Haematol. 52, 259-267.
- 11. Ratnoff, O.D. (1984) J. Lab. Clin. Med. 103, 653-659.
- 12. Bloom, A.L. (1984) Lancet i, 1452-1454.
- 13. Wood, I.W., Capon, O.J., Simonsen, C.C., Eaton, D.L., Gitschier, J., Keyl, B., Seeburg, P.M., Smith, D.H., Hollingshead, P., Wion, K.L., Delwart, E., Tuddenham, E.G.D., Vehar, G.A., and Lawn, R.M. (1984) Nature 312, 330-337.
- 14. Toole, J.J., Knopf, J.L., Wozney, J.M., Sultzman, L.A., Buecker, J.L., Pittman, D.D., Kaufman, R.J., Brown, E., Shoemaker, Ch., Orr, E.C., Amphlett, G.W., Foster, W.B., Coe, M.L., Knutson, G.J., Fass, D.N.,

and Hewick, R.M. (1984) Nature 312, 342-347. 15. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning Laboratory Manual. Cold Spring Harbor Laboratory. 16. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467. 17. Jaffe, E.A., Nachman, R.L., Becker, C.G., and Minick, C.R. (1973) J. Clin. Invest. 52, 2745-2756. 18. Willems, Ch., Astaldi, G.C.B., de Groot, Ph.G., Janssen, M.C., Gonsalves, M.D., Zeijlemaker, W.P., van Mourik, J.A., and van Aken, W.G. (1982) Exp. Cell Res. 139, 191-197. 19. Warner, J.R., Soeiro, R., and Birnboim, H.C. (1966) J. Mol. Biol. 19, 349-361. 20. Aviv, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408- 1412. 21. Rosen, J.M., Woo, S.L.C., Holder, J.W., Means, A.R., and O'Malley, B.W. (1975) Biochemistry 14, 69-78. 22. Wieslander, L. (1979) Anal. Biochem. 98, 305-309. 23. Gurdon, J.B., Lane, C.D., Woodland, H.R., and Marbaix, G. (1971) J. Mol. Biol. 233, 177-181. 24. Veerman, E.C.I., Stel, H.V., Huisman, J.G., and van Mourik, J.A. (1983) Thromb. Res. 33, 89-93. 25. Gubler, U., and Hoffman, B.J. (1983) Gene 25, 263-269. 26. Viera, J., and Messing, J. (1982) Gene 19, 259-268. Hanahan, D. (1983) J. Mol. Biol. 166, 557-580. 28. Helfman, D.M., Feramisco, J.R., Fiddes, J.C., Thomas, G.P., and Hughes, S.H. (1983) Proc. Natl. Acad. Sci. USA 80, 31-35. 29. Ruggeri, M.Z., and Zimmerman, T.S. (1981) Blood 57, 1140-1143. 30. Laemmli, U.K. (1970) Nature 227, 680-685. Towbin, H., Staehlin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350-4354. 32. Cabilly, S., Riggs, A.P., Pande, H., Shively, J.E., Holmes, W.E., Rey, M., Perry, L.J., Wetzel, R., and Heyneker, H.L. (1984) Proc. Natl. Acad. Sci. USA 81, 3273-3277. 33. Lehrach, H., Diamond, D., Wozney, J.M., and Boedtker, H. (1977) Biochemistry 16, 4743-4751. 34. Sodetz, J.M., Paulson, J.C., and McKee, P.A. (1979) J. Biol. Chem. 254, 10754-10760. 35. Stel, H.V., Sakariassen, K.S., Scholte, B.J., Veerman, E.C.I., van der Kwast, Th.H., de Groot, Ph.G., Sixma, J.J., and van Mourik, J.A. (1984) Blood 63, 1408-1415. 36. Reinders, J.H., de Groot, Ph.G., Gonsalves, M.D., Zandbergen, J., Loesberg, C., and van Mourik, J.A. (1984) Biochim. Biophys. Acta 804, 361-369. 37. Kefalides, N.A. (1981) Fed. Proc. 40, 327. 38. Edgell, C.-J.S., Reisner, H.M., and Graham, J.B. (1980) Brit. J. Haematol. 46, 613-620.

Note added in proof:

After submission of this manuscript, we learned that recently similar work has been reported (Lynch et al., 1985, Cell 41, 49-56). Their data and ours on the cloning of human vWF-cDNA are in agreement, except for some discrepancies in the nucleotide sequence of the 3'-untranslated region around position 262 (Fig. 6).