Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura

Xinglong Zheng‡§, Dominic Chung‡¶, Thomas K. Takayama‡, Elaine M. Majerus**, J. Evan Sadler**‡‡§§, and Kazuo Fujikawa¶

From the Departments of ‡Pathology, **Medicine, and §§Biochemistry and Molecular Biophysics and Howard Hughes Medical Institute, Washington University School of Medicine, St. Louis, Missouri 63110 and the Departments of ¶Pathology, University of Washington, Seattle, Washington 98195

Thrombotic thrombocytopenic purpura is associated with acquired or congenital deficiency of a plasma von Willebrand factor-cleaving protease (VWFCP). Based on partial amino acid sequence, VWFCP was identified recently as a new member of the ADAMTS family of metalloproteases and designated ADAMTS13. The 4.6-kilobase pair cDNA sequence for VWFCP has now been determined. By Northern blotting, full-length VWFCP mRNA was detected only in liver. VWFCP consists of 1427 amino acid residues and has a signal peptide, a short propeptide terminating in the sequence RQRR, a reopolyisin-like metalloprotease domain, a disintegrin-like domain, a thrombospondin-1 repeat, a Cys-rich domain, an ADAMTS spacer, seven additional thrombospondin-1 repeats, and two CUB domains. VWFCP apparently is made as a zymogen that requires proteolytic activation, possibly by furin intracellularly. Sites for Zn²⁺ and Ca²⁺ ions are conserved in the protease domain. The Cys-rich domain contains an RGDS sequence that could mediate integrin-dependent binding to platelets or other cells. Alternative splicing gives rise to at least seven potential variants that truncate the protein at different positions after the protease domain. Alternative splicing may have functional significance, producing proteins with distinct abilities to interact with cofactors, connective tissue, platelets, and von Willebrand factor.

Thrombotic thrombocytopenic purpura (TTP) is a syndrome characterized by microangiopathic hemolytic anemia and thrombocytopenia, and it may be accompanied by neurological dysfunction, renal failure, and fever (1, 2). TTP often strikes young adults, mainly females, suggesting an autoimmune etiology. If untreated, the mortality may exceed 90% (2), but plasma exchange therapy has reduced the mortality to less than 20% (3).

The basis for the efficacy of plasma exchange remains unknown, although a plausible model has been suggested in which the proteolysis of von Willebrand factor (VWF) plays a central role. In 1982, patients with chronic relapsing TTP were reported to have "unusually large" VWF multimers (UL-VWF) because of the absence of a VWF depolymerase activity. Binding of UL-VWF to platelets could promote microvascular thrombosis, platelet consumption, and hemolysis. Therapeutic infusion of plasma was proposed to replace the missing depolymerase activity or other factors and thereby limit VWF-dependent platelet thrombosis (4). Later discoveries have strengthened the relationship between VWF proteolysis and TTP. A plasma metalloprotease was identified that requires both calcium and zinc ions and cleaves the Tyr¹⁶⁰⁵–Met¹⁶⁰⁶ bond in the central A2 domain of the VWF subunit (5, 6). Cleavage was stimulated by shear forces like those occurring at sites of arterial thrombosis or by low concentrations of urea or guanidine. Furthermore, most adult patients with TTP were found to have congenital deficiency or an acquired autoantibody inhibitor of this VWF-cleaving protease (VWFCP) (7–9). These findings are consistent with the UL-VWF model of TTP pathogenesis. Alternatively, it is possible that failure to cleave another unknown substrate contributes to the pathogenesis of TTP.

VWFCP was recently purified from human plasma sources, and an N-terminal amino acid sequence was obtained (10, 11). This information showed it to be a member of the ADAMTS family of metalloproteases, named for the characteristic combination of a disintegrin-like and metalloprotease (reprolysin-type), with thrombospondin type 1 motif (12–14). Furthermore, the VWFCP gene was identified with the partially characterized C9ORF8 gene on chromosome 9 (11). Completion of the cDNA sequence now shows the VWFCP gene to be much more extensive than predicted. The deduced protein sequence indicates that VWFCP is similar in structure to other ADAMTS proteases but with unique features that suggest distinct modes of ligand binding. Alternatively spliced transcripts were characterized that change the domain architecture and may have functional consequences for the interaction of VWFCP with extracellular matrix, cell surfaces, substrates, and other proteins.

* This work was supported in part by National Institutes of Health Grant HL16919. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY055376.

§ Contributed equally.

‡ To whom correspondence should be addressed: Howard Hughes Medical Inst., Washington University School of Medicine, 660 S. Euclid Ave., Box 8022, St. Louis, MO 63110. Tel.: 314-362-9029; Fax: 314-454-0175; E-mail: esadler@im.wustl.edu.

† The abbreviations used are: TTP, thrombotic thrombocytopenic purpura; VWF, von Willebrand factor, VWFCP, von Willebrand factor-cleaving protease; UL, unusually large; kb, kilobase(s); PCR, polymerase chain reaction; nt, nucleotide(s); TSP1, thrombospondin-1.

Residues of VWF are numbered from the initiation codon, and Tyr¹⁶⁰⁵ corresponds to Tyr⁶⁴² of the mature VWF subunit sequence.
**EXPERIMENTAL PROCEDURES**

**Characterization of cDNA Clones**—Genomic DNA sequence containing the C9ORF8 locus was analyzed with the programs GENESCAN (15) and GraiIEXP (16) using the Oak Ridge National Laboratory website (combio.ornl.gov). This analysis identified a 37-kb potential gene that spanned C9ORF8 at the 5’ end (UniGene cluster Hs.149184) and the hypothetical protein DKPzP424C222 locus at the 3’ end (UniGene cluster Hs.131433). Expressed sequence tag clones from these clusters with IMAGE numbers 1925891 (GenBankTM AI346299), 1708760 (GenBankTM AI245927), 2462784 (GenBankTM AI927983), and 1654221 (GenBankTM AI023700) were sequenced with BigDye cycle DNA sequencing reagents (Applied Biosystems, Foster City, CA). DNA fragments between these sequences were generated by PCR with Marathon RACE-ready human liver cDNA and AdvantageTM GC 2 PCR kit (CLONTECH, Palo Alto, CA) using primers 5’-CCT CCG CCC AGA CAT TGG A-3’ and 5’-GGG CTG GCC AGA CAT TGG A-3’ (nt 830–2153 complement). The sequence between nt 830 and 2153 was amplified with primers 5’-GCT GCG GCC CCA GCG GAC ACG T-3’ and 5’-GGG CTG GCC AGA CAT TGG A-3’ (nt 3947–4529 complement). PCR products were cloned into plasmid pCR2.1 (Invitrogen, Carlsbad, CA) and sequenced. Membranes were washed sequentially with 2X SSC, 0.5% SDS at 50°, 3X SSC, 0.5% SSC, 0.1% SDS at 50°C, and 0.1X SSC, 0.5% SDS at 50°C for one h after 30 min prehybridization with ExpressHybTM (Stratagene). Hybridization signals were detected using [32P]dCTP-labeled human actin cDNA (23) and exposed for 1 h.

**RESULTS**

**Characterization of cDNA Clones**—The amino acid sequence of purified plasma VWFCP enabled the identification of the VWFCP gene with the hypothetical C9ORF8 gene on human chromosome 9. Sequence comparisons showed that VWFCP belongs to the ADAMTS family of metalloproteases (10, 11). The predicted 2.2-kb cDNA sequence of C9ORF8 contains several insertions and frameshifts and lacks all or parts of the signature propeptide, protease, and spacer domains of ADAMTS proteases (10, 11). Analysis of human chromosome 9 for predicted transcripts (15, 16) suggested that the C9ORF8 locus was part of a much larger transcription unit that contained at least 28 exons, spanned at least 37 kb, and included the adjacent hypothetical DKPzP424C222 locus (nt 4505–4529 complement). PCR products were cloned into plasmid pCR2.1 (Invitrogen, Carlsbad, CA) and sequenced.

**Characterization of cDNA Clones—** The amino acid sequence of purified plasma VWFCP enabled the identification of the VWFCP gene with the hypothetical C9ORF8 gene on human chromosome 9. Sequence comparisons showed that VWFCP belongs to the ADAMTS family of metalloproteases (10, 11). The predicted 2.2-kb cDNA sequence of C9ORF8 contains several insertions and frameshifts and lacks all or parts of the signature propeptide, protease, and spacer domains of ADAMTS proteases (10, 11). Analysis of human chromosome 9 for predicted transcripts (15, 16) suggested that the C9ORF8 locus was part of a much larger transcription unit that contained at least 28 exons, spanned at least 37 kb, and included the adjacent hypothetical DKPzP424C222 locus (nt 4505–4529 complement). PCR products were cloned into plasmid pCR2.1 (Invitrogen, Carlsbad, CA) and sequenced.

**Characterization of cDNA Clones—** The predicted amino acid sequence of human VWFCP contains 1427 amino acid residues (see Fig. 1 and Fig. 2). The predicted N-terminal signal peptide precedes a 41-residue propeptide that ends in the sequence QRRQ, which is followed by the experimentally determined N-terminal protein sequence of plasma VWFCP (10, 11). Similar basic propeptide cleavage sites occur in other ADAMTS proteases and are cleaved intracellularly by furin during biosynthesis (13). The next domains comprise the defining motifs of the ADAMTS family, a reprocollagen-like metalloprotease domain, a disintegrin-like domain, a thrombospondin-1 repeat (TSP1), and characteristic Cys-rich spacer domains (12-14). The Cys-rich domain of VWFCP has an RGDS sequence that could bind to integrin-like platelets or other cell types. An RGD sequence is present in the Cys-rich domain of ADAMTS-2 (procollagen I protease) (25) but not in the other known mammalian ADAMTS proteases. In the VWF spacer domain, the Ets domain is followed by seven more TSP1 repeats. ADAMTS proteases often have a variable number of TSP1 repeats in locations similar to those of VWFCP. For example, ADAMTS9 (long form) has 12 TSP1 repeats after its spacer domain (26, 27). With the exception of TSP1 repeat 5,
which is encoded by two exons, each of these additional TSP1 repeats in VWFCP is encoded by one exon (Fig. 2). TSP repeat 4 is unusual for the absence of two conserved Cys residues that are present, however, in the homologous TSP1 repeat 4 of mouse VWFCP (not shown). The TSP1 repeats of VWFCP are followed by two CUB domains (28), which were first identified in complement components C1r and C1s. VWFCP is the only known ADAMTS protease with CUB domains, and this distinctive domain structure is reflected in the evolutionary relationships among ADAMTS proteases. Comparison of the protease, TSP1, and Cys-rich and spacer domain sequences shows that VWFCP is the most divergent member of the family (Fig. 2).

The VWFCP protease domain has the expected hallmarks of the reprolysin (29) or adamalysin (30) family of metalloproteases. These include (Fig. 1) three His residues that coordinate the essential Zn$^{2+}$ ion in the sequence HE*XX - H*XX G*XX HD, the conserved Met 249 residue in a proposed Met turn, and residues Glu83, Asp173, Cys281, and Asp284, which are predicted to coordinate a structural Ca$^{2+}$ ion. The presence of both metal ion sites is consistent with the inhibition of enzyme activity by chelation of either zinc or calcium ions (5, 6). Sequence alignments and preliminary modeling based on the structure of adamalysin II (31) suggest that disulfide bonds connect Cys155–Cys208, Cys202–Cys281, and Cys242–Cys265.

Active VWFCP begins with amino acid residue Ala75 and has a calculated polypeptide mass of 145 kDa, compared with the apparent mass of 190 kDa for VWFCP purified from human plasma (11). The difference may be because of glycosylation. VWFCP has 10 potential N-glycosylation sites (Asn-Xaa-Thr/Ser) (Fig. 1), and the TSP1 repeats have additional sites for specific, unusual oligosaccharides (Fig. 3). TSP1 repeat 1 contains the sequence W*XX W, which often is modified in TSP1 repeats by attachment of a Glc-Man group to the C-2 atom of the first Trp (32). Seven of the eight TSP1 repeats also contain the conserved sequence CSX(S/T)CG, in which the hy-
Expression of VWFCP and Alternative Splicing—Northern blotting of human poly(A)⁺ mRNA identified a 4.7-kb mRNA in liver and a 2.4-kb mRNA in placenta and skeletal muscle (Fig. 4). Therefore, liver may be mainly responsible for the synthesis of full-length VWFCP. By the more sensitive method of reverse transcriptase-PCR using primers specific for the first CUB domain and 3' noncoding sequence, all tissues examined show PCR products characteristic of the VWFCP (not shown), suggesting that VWFCP or variant forms of it are expressed at low levels in all tissues. Accordingly, partial cDNAs for VWFCP were isolated from brain and prostate in addition to liver.

Comparison of the cDNA sequences with genomic sequences identified seven alternatively spliced transcripts, three of which were observed in liver. The use of alternate donor or acceptor sites gave rise to transcripts with deletions of nt 933–937 (liver), 2197–2212 (liver), 3997–4185 (liver), 4004–4185 (brain), 3510–3676 (prostate), or 2198–2528 (prostate). In another brain transcript, the use of an alternate acceptor site results in the insertion of 235 nt at the 5' end of exon 8. These insertions and deletions cause frameshifts and truncation after the metallocproteinase domain, the spacer domain, or the first CUB domain or cause an in-frame deletion of 56 amino acids between the eighth TSP1 domain and the first CUB domain.

DISCUSSION

The domain structure of VWFCP indicates that it belongs to the ADAMTS family of metalloproteases, and VWFCP has been designated ADAMTS13 (HUGO Gene Nomenclature Committee, www.gene.ucl.ac.uk/nomenclature). The properties of other members of this family may be relevant to understanding the biological function of VWFCP. ADAMTS1 is required for normal development of the genitourinary system; deficiency causes obstructive uropelvic fibrosis and defects in renal and adrenal morphology (33). Procollagen I N-proteinase (ADAMTS2) releases N-terminal propeptides from procollagens I and II, permitting the extracellular assembly of collagen fibrils. Deficiency of procollagen I N-proteinase causes human Ehlers-Danlos syndrome type VIIC, which is associated with fragile skin and hyperextensible joints (25). ADAMTS3 also has procollagen II N-proteinase activity (34). Aggrecanase-1 (ADAMTS4) (35–37), aggrecanase-2 (ADAMTS5) (38), and ADAMTS1 (37, 39) cleave proteoglycans of cartilage (aggrecan), brain (brevican), or articular wall (versican V1) in vitro at peptide bonds that also are cleaved in vivo. VWF may be thought of as a mobile connective tissue component that is incorporated at sites of vascular injury, where it mediates platelet adhesion and is cleaved by VWFCP (ADAMTS13). From this perspective, the known activities of ADAMTS proteases involve the processing of adhesive glycoproteins in extracellular matrix.

Other ADAMTS proteases have propeptides of ~160 to ~230 amino acid residues that contain two or three Cys residues. In many metalloproteases including the ADAMTS family, a Cys residue in the propeptide is proposed to maintain latency of the protease activity by coordinating the catalytic Zn²⁺ residue (40, 41). However, VWFCP has a propeptide of only 41 residues with one Cys that is not in a typical “cysteine switch” environment (Fig. 1). Whether a cysteine switch mechanism operates in human VWFCP is not known, and residue Cys⁵⁵ is not conserved in the propeptide of mouse VWFCP, being replaced by Leu (not shown). The VWFCP propeptide ends with the sequence RQR, which suggests that furin or a related processing protease cleaves it during synthesis. Other ADAMTS propeptides are cleaved at similar sites by furin, either in the Golgi or at the cell surface (42–44). Therefore, VWFCP may be secreted as an active protease. Alternatively, extracellular cleavage and activation of proVWFCP could be regulated, providing an additional level of biological control.

Structural domains that follow the disintegrin-like domain can determine the tissue localization of certain ADAMTS proteases and may function similarly in VWFCP. Recombinant ADAMTS1 binds the extracellular matrix of cultured cells. Matrix binding can be mediated by the spacer domain or by several TSP1 repeats and is reversed by heparin, suggesting that ADAMTS1 binds glycosaminoglycans (45). Recombinant ADAMTS12 also binds to extracellular matrix (44). In another context, the TSP1 repeats of platelet thrombospondin appear to bind several glycoproteins and proteoglycans on cell surfaces and in extracellular matrix (46), as well as fibrinogen or fibrin (47). TSP1 repeats were discovered recently to bear two unusual post-translational modifications, an α-mannosyl group on the C-2 position of Trp and the disaccharide Glc-Fuc-O-Ser/Thr in the motif CSX(S/T)CG (32). These modifications are located within segments of TSP1 repeats that are implicated in ligand binding (reviewed in Ref. 46). Consensus sites for C-mannosylation and O-fucosylation occur in the TSP1 repeats of VWFCP (Fig. 3) and may contribute to normal biological function.

The CUB domains of VWFCP are not found in other known ADAMTS proteases but are characteristic of a distantly related metalloprotease subfamily that includes procollagen C-proteinase (synonyms, mTollloid and BMP-1), mammalian tollloid-like proteins 1 and 2, Xenopus hatching enzyme UVS2, and sea urchin BP10/SpAN (reviewed in Ref. 48). All of these proteins contain an astacin-like metalloprotease domain followed by one or more epidermal growth factor domains and up to five CUB domains. The CUB domains of procollagen C-proteinase bind to procollagen, thereby contributing to substrate recognition (48). In addition, normal procollagen C-processing requires a cofactor protein, procollagen C-proteinase enhancer, that has two CUB domains and also binds procollagen (49). These examples suggest the CUB domains of VWFCP may interact with physiological ligands.

Alternative splicing is a common feature of ADAMTS gene expression (12–14), and seven variant transcripts for VWFCP were identified in this study that truncate the predicted protein sequence after the metalloprotease domain, the spacer domain, or the first CUB domain. Alternative splicing may account for the shorter 2.4-kb transcripts found in placenta and skeletal muscle, as well (Fig. 4). Whether these various transcripts can give rise to functional protein is not known. However, alternative splicing may explain some of the size heterogeneity of plasma VWFCP, which may consist of 150-, 140-, 130-, and 110-kDa species that share a common N-terminal sequence (10). Furthermore, alternative splicing could produce several protein products with distinct abilities to interact with cofactors, extracellular matrix, platelets, and VWF.

The association of VWFCP deficiency with TTP is convincing (7–9), but the mechanism of deficiency is apparent for only the one-half of adult patients with anti-VWFCP antibodies and the rare patients with congenital VWFCP deficiency. VWFCP cleaves VWF at a unique Tyr–Met bond (5, 6), and deficiency of VWFCP leads to the accumulation of UL-VWF multimers. However, the spectrum of physiological substrates for VWFCP may be broader than VWF alone, and the role of UL-VWF in the pathogenesis of TTP is not fully understood. The characterization of the VWFCP (ADAMTS13) cDNA will facilitate the investigation of these important biochemical and medical questions and may lead to the development of more specific treatments for TTP.
Acknowledgments—We thank David Ginsburg (University of Michigan) for helpful discussions and advice and Earl W. Davie for support and encouragement.

REFERENCES

Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura

Xinglong Zheng, Dominic Chung, Thomas K. Takayama, Elaine M. Majerus, J. Evan Sadler and Kazuo Fujikawa

doi: 10.1074/jbc.C100515200 originally published online September 13, 2001

Access the most updated version of this article at doi: 10.1074/jbc.C100515200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 49 references, 23 of which can be accessed free at http://www.jbc.org/content/276/44/41059.full.html#ref-list-1