REDOX DEPENDENT DISULFIDE BONDING OF VON WILLEBRAND FACTOR A1A2A3 DOMAINS

REDOX DEPENDENT DISULFIDE BONDING OF VON WILLEBRAND FACTOR A1A2A2 DOMAINS

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Cell and Molecular Biology

By

Esra Seyran
Ege University
Bachelor of Science in Life Sciences, 1999
University of Arkansas
Master of Science in Plant Pathology, 2004

May 2011
University of Arkansas

ABSTRACT

Von Willebrand factor (VWF) is a large glycoprotein that plays an important role in primary hemostasis by forming a bridge between subendothelial matrix and platelets. The most striking feature of VWF is its ability to dramatically shift from a non-adhesive plasma soluble protein to a multi-functional adhesive protein, as an activated surface, attractive to the flowing platelets. The segmental structure of VWF forms the basis of its multifunctional properties. Binding sites that are independent of multimer assembly but important for the hemostatic function are located in the A1A2A3 domains of VWF. VWF A1A2A3 domain contains six cysteine residues in three disulfides, one being a vicinal disulfide. Established view is that disulfide bonds are inert motifs, stabilizing the protein structure. However vicinal disulfides in proteins are reported to function as redox activated conformational switches. In order to understand the redox dependent disulfide bonding of VWF A1A2A3 domain, we conducted experiments in reduced and oxidized glutathione environment. Disulfide mapping was performed by digesting A1A2A3 domain using GluC and analyzing it by using matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS). It was evident from the MALDI-MS data that the changes occur in disulfide bonding pattern in the glutathione environment. The regulation of VWF structure and function is important in the context of cardiovascular diseases since they are associated with elevated VWF activity. Improvement of our understanding over the structure, function and conformation of VWF could possibly make this protein a drug target, advancements in this area is an interest to the pharmaceutical industry.

This dissertation is approved for
Recommendation to the
Graduate Council

Dissertation Director:

_____

Dr. Wesley Stites

Dissertation Committee:

_____

Dr. Roger Koeppe

_____

Dr. Joshua Sakon

_____

Dr. Robert Beitle

DISSERTATION DUPLICATION RELEASE

I hereby authorize the University of Arkansas Libraries to duplicate this dissertation when needed for research and/or scholarship.

Agreed

_____

Esra Seyran

Refused

_____

Esra Seyran

DEDICATION

I dedicate this dissertation to my

aunt Naile Suna Ilterberk

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
OVERVIEW OF VON WILLEBRAND FACTOR

ABSTRACT

Haemostasis is a delicately balanced process to prevent blood loss during vascular injury. Haemostasis involves a complex network of plasma proteins and coagulation factors. A key player in haemostasis is von Willebrand factor (VWF). VWF is a multimeric, glycoprotein functions as a bridge between subendothelial structures at sites of vascular injury and blood platelets, apart from this role it also stabilize and protect coagulation factor VIII. My Ph.D. dissertation examines the redox dependent disulfide bonding of von Willebrand factor. During the course of our research efforts, we performed many of the experiments with a protein isolated from the human blood; we have developed a novel method to map the disulfide bonds in proteins using mass spectrometry. We examined disulfide bonding pattern of hemostaticaly functional A1A2A3 domains in both reduced and oxidized glutathione environments. We observed novel disulfide bond rearrangements between the A1 and A2 domain. The regulation of VWF structure and function is important in the context of cardiovascular diseases since they are associated with elevated VWF activity. Improvement of our understanding over the structure, function and conformation of VWF could possibly make this protein a drug target, advancements in this area is an interest to the pharmaceutical industry.

INTRODUCTION

von Willebrand factor (VWF) is a large glycoprotein that plays an important role in primary hemostasis by forming a bridge between subendothelial matrix and platelets (*1*) . VWF has two major hemostatic functions, mediating the adhesion of platelets to subendothelium at sites of vascular injury and stabilization and transport of coagulation factor VIII (*2*). The most striking feature of VWF is its ability to dramatically shift from a non-adhesive plasma soluble protein to a multi-functional adhesive protein, making an activated surface, attractive to the flowing platelets. VWF does not spontaneously bind to platelets in the absence of vascular injury however, injury to the endothelial layer lead to the exposure of subendothelial connective tissue and initiates the events of VWF mediated thrombogenesis. Initially, VWF binds to the components of subendothelium which results in a conformation change that enable the interaction of VWF with GPIb-IX-V complex, thereby the initiation of platelet adhesion (*3*). Under high shear stress conditions of blood flow VWF is essential for normal platelet adhesion and thrombus formation but mediation of platelet adhesion is not exclusive to VWF. Depending on the local hemodynamic conditions other adhesive proteins such as fibrinogen can assume these functions, however shear stress induced conformational changes of VWF may be critical for the regulation of thrombogenesis (*4*).

The second major hemostatic function of VWF is to transport coagulation factor VIII and localize it to the vessel wall upon vascular injury. VWF has been reported to prevent degradation and premature inactivation of factor VIII by the activated protein C (*5*) . Once secreted into plasma, factor VIII binds to VWF. Upon initiation of coagulation thrombin activates factor VIII and VWF binding site get disrupted then factor VIII

participates in coagulation cascade where it serves as the cofactor for factor IXa in the activation of factor X. Genetic deficiency of factor VIII is the cause of hemophilia A. The factor VIII binding site on VWF has been identified between residues 764-1035 of mature VWF in the D'D3 domains. Mutations that disrupt this binding site result in decreased binding capacity for factor VIII and cause autosomal recessive form of von Willebrand disease that resembles mild form of hemophilia (*6*),(*7*),(*8*). It has been reported that under shear conditions Factor VIII binding
accelerate proteolytic cleavage of VWF by ADAMTS-13 (*9*).

The plasma concentration of VWF is 10µg/ml, 4-20 µg/ml is considered the normal range (*10*). Elevated plasma VWF levels are reported in atherosclerosis, cancer, diabetes, liver disease and renal failure. One of the major causes of morbidity in developed countries is arterialthrombosis caused by rupture of an atherosclerotic plaque which causes excessive platelet
adhesion and thrombus formation resulting in diseases such as myocardial infarction and stroke. The plasma concentration of VWF is an established predictor of cardiovascular disease (*11*) . Elevated levels and abnormal processing of VWF are associated with isochemic heart disease, thrombocytopenic purpurea, arterial thrombosis (*12, 13*), peripheral-pulmonary vascular disease, cerebrovascular disease (*13*).

Von Willebrand disease (VWD) is the most common bleeding disorder in humans . Affected individuals reflect platelet dysfunction, excessive bleeding and hematomas symptoms associated with defective factor VIII or platelet binding functions of VWF. VWD is characterized by qualitative; type one and type three and quantitative; type two abnormalities of VWF. The prevalence of type one is estimated about 1% of the general

population, accounting for more than 70% of VWD. The inheritance is autosomal dominant . Type two represents 10-15% of VWD cases and causes moderate to mild bleeding. Increase in proteolytic fragments corresponding to 176kDa and 140kDa is common in type 2 patients due to their increased susceptibility to ADAMTS-13 cleavage. Type three VWD is characterized by severe bleeding less than 10% factor VIII activity and severely decreased VWF levels (*14*) .

The gene for VWF is on chromosome 12, over 178 kilobases contains 52 exons (*15*) . The mRNA transcript is translated to produce a 2813 amino acid pre-pro-protein in the endoplasmic reticulum of endothelial cells. The cDNA of translation product has approximately 360 kDa molecular weight. The pre-proprotein contains a 22 residue signal peptide, a propetide with 95 kDa and a mature subunit with 260 kDa molecular weight. The 22 amino acid signal peptide directs VWF protein to the endoplasmic reticulum, 741 amino acid propeptide is required for multimerization and for targeting VWF to storage granules. Propeptide cleavage occurs in the Golgi organelle by the furin enzyme, recognizes basic Lysine-Arginine peptide bond located at 762-763 position (*16*). The mature protein is 2050 amino acids and contains all the functional domains. VWF undergoes dimerization in the endoplasmic reticulum through disulfide bond formation between cysteine residues located between the carboxyl terminus of the mature subunit . The endoplasmic reticulum provides optimum conditions for the spontaneous oxidation of sulfydryl groups and for the formation of disulfide bonds, then VWF dimers are transported to the Golgi organelle, where the acidic environment in is more conducive for the multimer formation. Multimerization of VWF involves the formation of interchain

disulfide bonds between cysteine residues in the D3 domains of two VWF dimers (*17*), (*18*), (*19*), (*20*).

Disulfide bonds of VWF been identified after partial proteolysis by direct observation of phenylthiohydantoin of cystines during the Edman degradation of isolated peptides. Fifty-two half cystine residues have been paired in this experiment (Note: amino acids are numbered starting from the initial methionine residue in the pre-pro VWF and single amino acid coding is used in accordance with the most recent ISTH VWF SSC nomenclature guidelines; www.vwf.group.shef.ac.uk/index.html) (Table 1.1). Additional clusters of disulfide bonds were reported to be evident. Intersubunit disulfide bonds reported to be localized in an interior region (A domain) and a carboxyl-terminal region (CK). Disulfide bonds appeared to established local (within the same domain) with less than 170 residues rather than more distant parts of the same polypeptide chain (*19*).

Table 1.1 Described disulfide bonds of VWF.

Disulfide Bond	Distance	Domain
Cys767 ↔Cys 808	42	D2
Cys776 ↔Cys804	29	D2
Cys810 ↔Cys821	12	D2
Cys867 ↔Cys 996	130	D3
Cys889 ↔Cys1031	143	D3
Cys898 ↔Cys 993	96	D3
Cys914 ↔Cys 921	8	D3
Cys1060 ↔Cys 1084	25	D3
Cys1071 ↔Cys 1111	41	D3
Cys1089 ↔Cys 1091	3	D3
Cys1126 ↔Cys 1130	5	D3
Cys1149 ↔Cys 1169	21	D3
Cys1153 ↔Cys 1165	13	D3
Cys1196 ↔Cys 1199	4	D3
Cys1234 ↔Cys 1237	4	D3
Cys1272 ↔Cys 1458	187	A1
Cys1669 ↔Cys1670	2	A2
Cys1686 ↔Cys 1872	187	A3
Cys1879 ↔Cys1904	26	D4
Cys1899 ↔Cys 1940	42	D4
Cys1927 ↔Cys 2088	162	D4
Cys1950 ↔Cys 2085	136	D4
Cys1972 ↔Cys 2123	152	D4
Cys1993 ↔Cys 2001	9	D4
Cys2724 ↔Cys 2774	51	CK
Cys2739 ↔Cys 2788	50	CK
Cys2750 ↔Cys 2804	55	CK
Cys2754 ↔Cys 2806	53	CK

In the endothelial cells, the majority of newly synthesized VWF multimers is secreted constitutively, whereas the remaining, approximately 5% is stored inside the Weibel-Palade bodies, that are rod-shaped organelles specific to the endothelial cells. Endothelial VWF is secreted both to the lumen and the extracellular subendothelial VWF is released from the platelet α granules after stimulation. Megakaryocytes also synthesize

VWF and store it in α-granules, variety of secretagogues such as thrombin, histamine, fibrin, ephineprine and forskolin enable the release of VWF from storage granules (*21*). Once secreted into the plasma, circulating VWF has a half-life of 8 to 12 hours. VWF circulates in plasma as a series of variable size multimers, ranging between 500 to 200,000 kDa. Regulation of multimer size is crucial for hemostasis. In the absence of high molecular weight multimers functional activity of VWF is decreased in contrary arterial thrombosis may occur in the presence of ultra large VWF multimers (*22*).

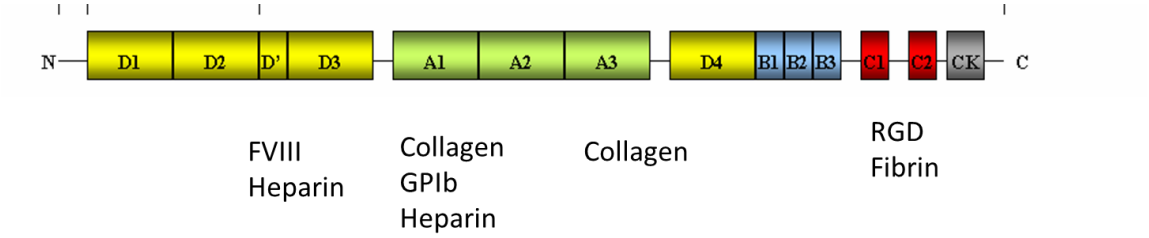
The size of circulating VWF multimers are regulated by plasma protease ADAMTS-13, is a member of ADAMTS family (a-disintegrin and metalloprotease with thrombospondin-1 repeats). Following cleavge VWF multimers acquire a globular shape preventing further proteolysis. ADAMTS13 proteolysis produces fragments in 140kDa and 176kDa size (*23*) . Beside the established role of ADAMTS-13 in the regulation of VWF size, number of studies have proposed that the effective size of VWF may be regulated by the self-association of the protein (*24*), (*25*) . Ulrichts, Vanhoorelbeke et al. (*26*) suggests that multiple domains in VWF may participate in the self-association process , according to Choi, Aboulfatova et al.(*27*) , this process may be accompanied by novel disulphide bond formation in the protein, that shear stress between 2000-5000/s changed the protein conformation and allowed residues in the D3 and C domains of VWF (Cys889, Cys 898, Cys2448, Cys2451, Cys 2490, Cys2491, Cys2453, Cys2528, and Cys2533) to undergo shear induced disulfide bond formation. It has been demonstrated that thrombospondin-1 (TSP-1) also directly controls VWF multimer size by reducing the disulfide bonds. This property has been attributed to the C-terminal calcium binding moiety within the TSP-1 molecule and requires a free thiol at position 974.(*28*) .

Bonnefoy et al. (*29*) recently described that TSP-1 rather protects subendothelium-bound VWF from degradation by the ADAMTS13. Proteases such as calpains, cathepsin G and elastases can also reduce the size of VWF with nonspecific cleavages (*30*).

VWF is highly glycosylated, 18.7 % its molecular weight is attributed to carbohydrates . In the endoplasmic reticulum early processing of carbohydrate chains occurs and twelve N- linked carbohydrates are added to the protein. Then VWF dimers are transported to the Golgi where ten O-linked carbohydrates are added. The carbohydrate structures of VWF has function in multimer formation, binding to collagen , protection against proteolysis and binding to platelet GPIb (*31*) .

VWF has four distinct homology domains (A-D). VWF contains binding sites for its ligands within a repeated multi-domain structure; D1-D2-D'-D3-A1-A2-A3-D4-B1-B2-B3-C1-C2-CK (Figure 1.1) . The segmental structure of VWF forms the basis of its multifunctional properties. The separate domains within the VWF subunit exhibit specific interactions with other molecules.

Figure 1.1 Domain structure of VWF.



In the A1 and A3 domains of VWF, collagen binding sites have been identified. VWF lacking the A1 domain is demonstrated to bind collagen type III and in invitro binding assays deletion of A3 domain prevented VWF binding to collagen. The results of this study showed that the A3 domain of VWF contains the primary binding site for collagen. A1 domain of multimeric VWF bind to platelet glycoprotein GPIb. The binding of VWF to sulfated glycolipids involves residues 1332 to 1347 and 1391 and 1409 of the A1 domain. GPIb binding to VWF A1domain, has shown to be induced by interaction with ristocetin and botrocetin. Indeed, the unchallenged VWF cannot bind to GPIb receptor hence in vitro static studies always require non-physiologic modulators, which can induce conformational changes comparable to those in vivo. Ristocetin is an antibiotic, its interaction induce a conformational change of the A1 domain which allows the interaction with GPIb by charge neutralization. Botrocetin , a snake venom protein , binds to positively charged residues in the VWF A1 domain, thereby inducing GPIb interaction. Although these interactions do not reflect biological functions, ristocetin and botrocetin binding found to be mimicking the shear-induced VWF binding to GPIb (*32*).

A1 domain has binding affinity for platelets GpIb-IX-V complex, binds to subendothelial VWF with the N-terminal portion of the GPIb. The GpIb-IX-V is a large

receptor complex of the leucine-rich repeat family (LRR). The binding site for this interaction in VWF is not functional under static conditions but active under elevated shear stress (*33*) . The binding site for sulfatides were reported to be located within residues 1275-1436 in the A1 domain loop . The Major heparin binding site on VWF resides in A1 domain between residues 1328 and 1341 (*34*) .The A2 domain does not have a known adhesive function, and deletion of this domain from VWF does not significantly affect either collagen or GPIb binding . However, A2 contains a proteolytic cleavage site for the plasma protease ADAMTS13 (*35*) . The size of VWF is regulated by the plasma protease ADAMTS13 that recognizes the peptide bond Y1605-M1606 in the A2 domain of VWF. The ultra large VWF multimers are released from the endothelial cells attaches to the cell surface via P-selectin and the shear force of the flowing blood helps expose the ADAMTS binding sites in the A1 and A3 domains and cleavage site in the A2 domain (*36*) . Mutations localized in the A2 domain of the VWF facilitate two distinct pathogenic mechanisms; mutations that are characterized by impaired intracellular transport secretion and storage of high molecular weight multimers or mutations that are characterized by increased susceptibility of the VWF protein to proteolysis by the ADAMTS13 metalloproteinase. A number of mutations have been identified located near this site, such as L1503Q, which lead to enhanced proteolysis of the VWF protein. Based on a molecular dynamics simulation on a homology model of the A2 domain it has been shown that so-called group I mutations cause significant deviations of the wild type structure in multiple regions of the VWF protein, whereas group II mutations cause single loop displacements near the proteolysis site which may affect interactions between ADAMTS13 and VWF. Oxidized and unoxidized VWF

peptides were incubated with recombinant ADAMTS13, substrate cleaved the oxidized VWF peptide much more slowly than unoxidized peptide (*37*) . A different study demonstrated that peroxynitrite-treated VWF multimers are more resistant to proteolysis by ADAMTS-13 and inhibitory effect of this substance increases with oxidant concentrations (*38*). The common finding of both studies is that oxidation of the VWF substrate causes a significant decrease in the ADAMTS-13 catalytic specificity. These effects in vivo may lead to a qualitative and quantitative alteration of circulating VWF. According to proposed mechanism stretching force of multimeric VWF needs to exceed 390pN for holding platelets at shear stress of 20dyne/cm2 this is likely to cause conformational changes which can expose VWF A2 domain (*39*). Physiological fluid shear up to 2000/s for 12s in a capillary tube has been shown to augment VWF A2-domain proteolysis by ADAMTS-13 (*40*) in the same extent with the presence of 1.5M GdHCl, has also shown to increase the ADAMTS-13 proteolysis of VWF (*41*). Activity of ADAMTS13 is difficult to assess in vitro without partial denaturation of the substrate. It appears obvious that extreme conditions are necessary to allow for VWF cleavage in vitro. Furlan et al. (*35*) used 1.0 M of urea, Ba2+ ions and a low ionic strength buffer, Tsai et al (*40*) denatured the substrate using guanidine hydrochloride, with no further addition of non-physiologic salts. These data demonstrated that VWF needs to be conformationally challenged to expose its A2 domain in order to let the ADAMTS13 attack over Tyrosine1605-Methione1606 bond. The assays used before 2005 were all employed denaturing conditions developed by Tsai et al. and Furlan et al., In 2005, Kokame et al. (*42*), presented a FRET study that does not require denaturation, a more physiologic approach. The recombinant VWF peptide was covering the residues D1596-

R1668, however this assay is unlikely fully simulate the in vivo situation due to the truncated nature of the substrate.

The C-terminal end of protein VWF contains B and C domains. VWF contains a RDG sequence in its C1 domain between 2507-2510 residues. The RGD sequence in VWF is the only known binding site for platelet integrin GP IIb/IIIa. Integrin GpIIbIIIa binds several ligands other than VWF due to its binding affinity for the RGD sequence which includes fibronectin, thrombospondin and vitronectin. (*43*) . The C1 and C2 domains have structural homology with collagen and thrombospondin. The C1 and C2 domains contain a binding site for fibrin . VWF forms disulfide linked dimers in endoplasmic reticulum in a tail to tail orientation between adjacent cysteine knot domains (CK) of monomers. VWF contains a 94 amino acid long cysteine knot domain, at its C-terminal and it is structurally homologous to the CK domains of norrin proteins. Interchain disulfide bonds in the CK domain has determined after cyanogen bromide digestion, using partial reduction and alkylation followed by reverse phase HPLC and mass spectrometry analysis. When cyanogen bromide cleaved peptides were analyzed C2771, C2773 and C2811 residues were identified as the residues involved in interchain disulfide bonds. The importance of CK domain for VWF dimerization confirmed in a study where the expression of recombinant, CK domains with C2771A, C2773A AND 2881A mutations resulted in defective dimerization and secretion of monomers instead of dimers (*20*).

The D domain contain 52 cystein residues and have homology with the D domains of murine proteins . The binding site for factor VIII lies in the D domain between the residues 763-1035. The D1 and D2 domains comprise the VWF propeptide

and a portion of the N-terminal of the mature protein. D1 and D2 domains are required for multimerization, Detailed analysis of the secondary structure of VWF revealed that inter-molecular disulfide bonds are localized in an amino-terminal region in D1D2 domains and carboxy-terminal region , beyond the domain C2 (*44*) . Each D1 and D2 domains contain one CXXC motifs which is the signature of the active site of disulfide isomerase enzymes. When the vicinal cysteine residues were distrupted by the insertion of a glycine residue by changing the sequence to CXGXC, VWF multimerization was inhibited. Many of the cysteine residues in the D3 domain have been shown to involve in intrachain disulfide bonds that are important for the structural integrity of the protein . Two heparin binding sites between the residues 1212-1491 and 764-1035 are located in the D' and D3 domain. When amino acids C1272, V1314, R1315 and R1374 mutated decreased heparing binding observed . Removal of D'D3 domain resulted in increased clearance of plasma VWF whereas removal of D4 domain resulted in decreased plasma clearance (*45*).

Studies using electron microscopy, shown VWF to be a flexible, elongated molecule with small nodules at irregular intervals, with a width of 2.5 nm and a maximum extended length of 120nm. Studies using quasi electric light scattering have shown that the predominant structure is a loosely coiled "ball of yarn" with a diameter 100-300nm, and proved the patch like globular appearance of secreted VWF associated with the surface of platelets (*46*). When atomic force microscopy used to image VWF in physiological buffer solution absorbed on hydrophobic monolayer molecule, VWF observed as globular, major axis diameter was 106-149 nm and minor axis diameter was 77-81nm, it appeared to be composed of overlapping globular domains, similar to the

"ball of yarn" described in electron microscopy studies. Globular domains were classified as large (L: 66nm major axis diameter) or small (S: 38nm major axis diameter) and were shown to align as L-S-L-L-S-L (*18*). Due to the colossal size of the molecule, X-ray crystallography studies have been conducted only on isolated recombinant A1,A2 A3 domains and have allowed structural characterization of individual protein domains either alone (*47*) or when in complex with antibodies (*48*) , botrocetin (*49*) or GpIb receptor (*50*) . The characterization of the crystal structure has described the A1 domain with an overall cuboidal shape, which is characterized by 5 parallel β strands and 1 antiparallel β strand which are centrally located, and flanked on either side by 3α helices. There are 6 flat faces, including a mostly basic and a mostly acidic one. A salt bridge network runs around the Al domain. Molecular homology modeling studies of the A1 domain suggested that the a crevice formed close to the side at which β strands wind in opposite directions represents an active site for ristocetin binding. Celikel et al., using X-ray crystallography reported the pairing of A1 domain residues close to the crystallographic 2-fold axis of symmetry, with salt bridges occurring between pairs of molecules.

Figure 1.2. VWF A1 domain. From PDB file 1M10 (*50*).

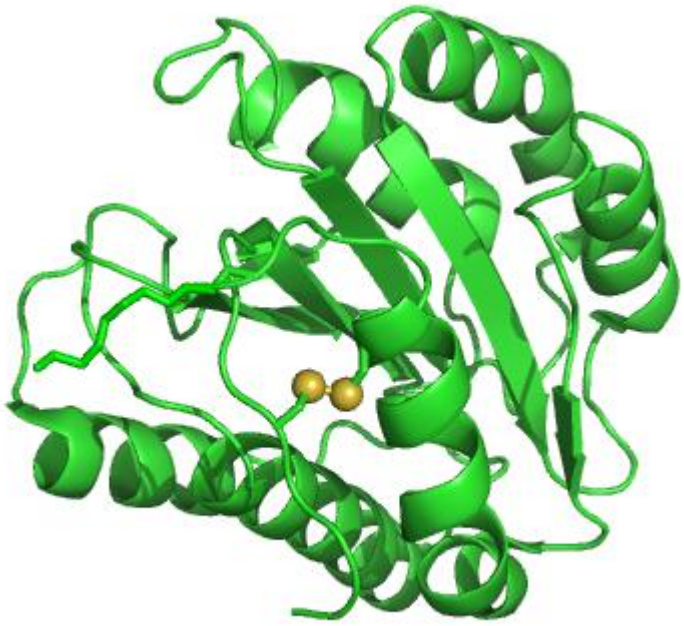
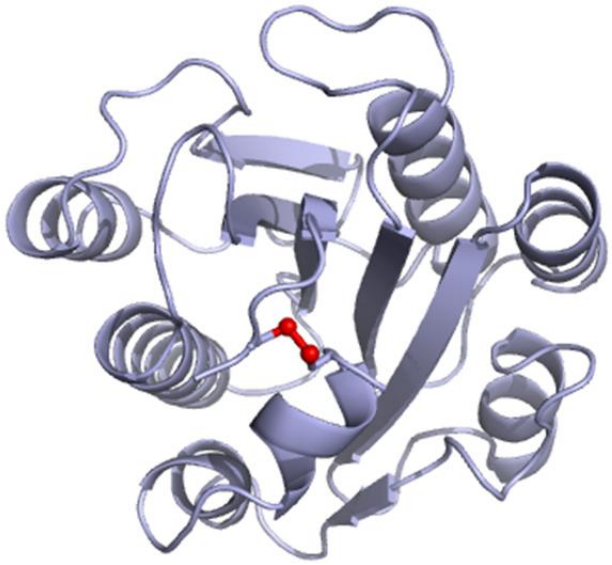


Figure 1.3. VWF A3 domain. From PDB file 1ATZ (*51*) .



According to the ribbon structure of the VWF A1 (Figure 1.2) and the VWF A3(Figure 1. 3) domain , The structural similarity of two domains can be appreciated as both A1 and A3 exist of six central beta sheets, surrounded by seven alpha helices. Both A1 and A3 contain a disulfide bond creating a loop, which is not the case in domain the A2 domain. (Figure 1.4).

Figure 1.4. VWF A2 domain. From PDB file 3GXB (52) .

Crystallographic studies revealed that both A1 and A3, adopt a typical α/β fold consisting of a central hydrophobic β-sheet neighbored on each side by amphipathic α-helices. This fold has also been deteceted in the homologous integrin I domains , in contrast with the I domains, metal ion dependent adhesion site motif is absent in the structures of A1 and A3 domain.

1.9-Å crystal structure of the A2 domain has been recently solved (*52*) . Crystal structure revealed a proof of evolutionary adaptations to the shear sensor function of A2 domain. VWF A2 contains a loop surrounded by 4 α helix, and a cis-proline. The central β strands are loosely packed The ADAMTS-13 cleavage site is buried in the central

hydrophobic core. Two vicinal Cys1669-Cys1670 residues, that are linked by a vicinal disulfide bond are shown to be buried in a hydrophobic pocket.

The x-ray crystal structure of A1 domain (PDB 1M10) covers 208 residues and defines the sequence from residue 1268 to 1476. The x-ray crystal structure of A2 domain (PDB 3GXB) covers 184 residues and defines the sequence from residue 1495 to 1673. The x-ray crystal structure of A3 domain (PDB 1ATZ) covers 189 residues and defines the sequence from residue 1688 to 1877. Between the A1 and A2 domains 19 residues are not covered in the crystal structures and between the A2 and A3 domains 15 residues are not covered in the crystal structures.

In a recent study VWF A2 domain with ΔΔCys1669Cys1670 and VWF A2 domain with vicinal Cys1669Cys1670 were compared in ADAMTS13 proteolysis and circular dichorsim experiments. The results of experiments demonstrated A2 domain with native vicinal cysteins was resistant to proteolysis and thermal unfolding, mutagenesis of Cys1669 Cys1670 enhanced unfolding and increased susceptibility to ADAMTS13 proteolysis (*53*).

Hematological consequences of the conformational transitions at the domain level remains to be unknown. The disulfide loop between the residues Cys1272-Cys1458 in A1 domain appears to manage the conformational transitions that affect the affinity of VWF for GPIb whereas the disulfide loop between the residues Cys1686 to Cys1872 in A3 domain likely plays a role in collagen binding. The binding of VWF to platelet receptor GpIb results in intermolecular bonds with high tensile strength (*22*), (*24*), (*54*). This molecular interaction allows platelet capture at sites of vascular injury under high fluid shear conditions. The binding of plasma VWF to platelet receptor GpIb leads to platelet

activation (*33*). Based on molecular dynamics simulation models of Yago et al. (*55*) at lower forces, the salt bridge between D1269 and R1306/R1450 stabilizes the VWF A1domain orientation at 75° to horizontal, at intermediate forces salt bridge between D1269 and R1306/R1450 is ruptured. This rupture of salt bridge extends the A1 N-terminal sequence which results in a torque to rotate the A1 to 64° to horizontal. This rotation helps β finger to slide over the top surface of A1, during which GPIb E14 forms a salt bridge with A1 R1334 (*55*).

In the absence of shear stress, atomic force microscopy resolution of the VWF molecule shows a globular condensed three dimensional structure. Exposure of VWF to a shear stress induces a conformational transformation from a globular state to an extended chain conformation with the exposure of intramolecular domains (*56*) . According to several studies VWF undergoes structural alterations under fluid shear. It has been shown that physiological fluid shear up to 2000/s for 12s in a capillary tube augment VWF A2domain proteolysis by ADAMTS13 (*40*) . In a different study shear rates in range of 2000-5000/s cause novel disulphide bond formation in the protein (*27*), Shankaran, Alexandridis et al. in 2003 demonstrated the self-association of VWF at shear rates of 2300-6000/s (*25*). It has been proposed that VWF undergoes conformational changes in order to expose the hidden GPIbα binding site for effective binding to platelets. When the disulfide bond of the A1 domain was reduced, and intermediate structures was observed via circular dichroism and fluorescence at low shear stress enhanced binding to platelets was observed, at higher shear binding capacity to platelets was significantly diminished whereas the disulfide intact A1 domain firmly adhered to platelets most effectively at high shear (*56*). These results imply that the wild

type A1 domain unfolds to an intermediate at high shear and bind to platelets as a result of this conformational transition. Recently Auton et. al., (*57*) have demonstrated that the A1 domain transition from native to reduced state increases the binding affinity for GPIbα.

Recent studies demonstrated specific interactions within the A domain. In this regard, Martin et al. showed that the VWF A2 domain interacts with the A1 domain, and this interaction may block VWF-GpIbα interactions in the native protein. Nishio et al. (*23*) suggested that the A1 domain hinders the action of the metalloprotease ADAMTS-13 on the VWF A2-domain. Binding interactions via the A1 domain enhance the rate of proteolysis mediated by ADAMTS-13. According to these studies changes at the domain-level may have important functional consequences.

REFERENCES

1. Ruggeri, Z. M., and Ware, J. (1992) THE STRUCTURE AND FUNCTION OF VON WILLEBRAND FACTOR, *Thrombosis and Haemostasis 67*, 594-599.

2. Weiss, H. J. (1991) VONWILLEBRAND-FACTOR AND PLATELET-FUNCTION, *Annals of the New York Academy of Sciences 614*, 125-137.

3. Mohri, H., Fujimura, Y., Shima, M., Yoshioka, A., Houghten, R. A., Ruggeri, Z. M., and Zimmerman, T. S. (1988) STRUCTURE OF THE VONWILLEBRAND-FACTOR DOMAIN INTERACTING WITH GLYCOPROTEIN IB, *Journal of Biological Chemistry 263*, 17901-17904.

4. Ikeda, Y., Handa, M., Kawano, K., Kamata, T., Murata, M., Araki, Y., Anbo, H., Kawai, Y., Watanabe, K., Itagaki, I., Sakai, K., and Ruggeri, Z. M. (1991) THE ROLE OF VONWILLEBRAND-FACTOR AND FIBRINOGEN IN PLATELET-AGGREGATION UNDER VARYING SHEAR-STRESS, *Journal of Clinical Investigation 87*, 1234-1240.

5. Koppelman, S. J., vanHoeij, M., Vink, T., Lankhof, H., Schiphorst, M. E., Damas, C., Vlot, A. J., Wise, R., Bouma, B. N., and Sixma, J. J. (1995) Requirements of von Willebrand factor to protect factor VIII from inactivation by activated protein C, *Blood 86*, 266-266.

6. Hamer, R. J., Koedam, J. A., Beeservisser, N. H., Bertina, R. M., Vanmourik, J. A., and Sixma, J. J. (1987) FACTOR-VIII BINDS TO VONWILLEBRAND-FACTOR VIA ITS MR-80000 LIGHT CHAIN, *European Journal of Biochemistry 166*, 37-43.

7. Wise, R. J., Dorner, A. J., Krane, M., Pittman, D. D., and Kaufman, R. J. (1991) THE ROLE OF VONWILLEBRAND-FACTOR MULTIMERS AND PROPEPTIDE CLEAVAGE IN BINDING AND STABILIZATION OF FACTOR-VIII, *Journal of Biological Chemistry 266*, 21948-21955.

8. Kao, K. J., Pizzo, S. V., and McKee, P. A. (1981) A RADIORECEPTOR ASSAY FOR QUANTITATING PLASMA FACTOR-VIII-VONWILLEBRAND PROTEIN, *Blood 57*, 579-585.

9. Cao, W. J., Krishnaswamy, S., Camire, R. M., Lenting, P. J., and Zheng, X. L. (2008) Factor VIII accelerates proteolytic cleavage of von Willebrand factor by ADAMTS13, *Proceedings of the National Academy of Sciences of the United States of America 105*, 7416-7421.

10. Sadler, J. E. (1998) Biochemistry and genetics of von Willebrand factor, *Annual Review of Biochemistry 67*, 395-424.

11. Blann, A. D., and McCollum, C. N. (1999) von Willebrand factor and soluble thrombomodulin as predictors of adverse events among subjects with peripheral or coronary atherosclerosis, *Blood Coagulation & Fibrinolysis 10*, 375-380.

12. Moake, J. L., Turner, N. A., Stathopoulos, N. A., Nolasco, L. H., and Hellums, J. D. (1986) INVOLVEMENT OF LARGE PLASMA VONWILLEBRAND-FACTOR (VWF) MULTIMERS AND UNUSUALLY LARGE VWF FORMS DERIVED FROM ENDOTHELIAL-CELLS IN SHEAR-STRESS INDUCED PLATELET-AGGREGATION, *Journal of Clinical Investigation 78*, 1456-1461.

13. Lip, G. Y. H., and Blann, A. D. (1995) VON-WILLEBRAND-FACTOR AND ITS RELEVANCE TO CARDIOVASCULAR DISORDERS, *British Heart Journal 74*, 580-583.

14. Rodeghiero, F., Castaman, G., and Dini, E. (1987) EPIDEMIOLOGIC INVESTIGATION OF THE PREVALENCE OF VONWILLEBRANDS DISEASE, *Blood 69*, 454-459.

15. Mancuso, D. J., Tuley, E. A., Westfield, L. A., Worrall, N. K., Sheltoninloes, B. B., Sorace, J. M., Alevy, Y. G., and Sadler, J. E. (1989) STRUCTURE OF THE GENE FOR HUMAN VONWILLEBRAND-FACTOR, *Journal of Biological Chemistry 264*, 19514-19527.

16. Rehemtulla, A., and Kaufman, R. J. (1992) PREFERRED SEQUENCE REQUIREMENTS FOR CLEAVAGE OF PRO-VONWILLEBRAND FACTOR BY PROPEPTIDE-PROCESSING ENZYMES, *Blood 79*, 2349-2355.

17. Wagner, D. D., Lawrence, S. O., Ohlssonwilhelm, B. M., Fay, P. J., and Marder, V. J. (1987) TOPOLOGY AND ORDER OF FORMATION OF INTERCHAIN DISULFIDE BONDS IN VONWILLEBRAND-FACTOR, *Blood 69*, 27-32.

18. Fowler, W. E., Fretto, L. J., Hamilton, K. K., Erickson, H. P., and McKee, P. A. (1985) SUBSTRUCTURE OF HUMAN VONWILLEBRAND-FACTOR, *Journal of Clinical Investigation 76*, 1491-1500.

19. Marti, T., Rosselet, S. J., Titani, K., and Walsh, K. A. (1987) IDENTIFICATION OF DISULFIDE-BRIDGED SUBSTRUCTURES WITHIN HUMAN VONWILLEBRAND-FACTOR, *Biochemistry 26*, 8099-8109.

20. Katsumi, A., Tuley, E. A., Bodo, I., and Sadler, J. E. (2000) Localization of disulfide bonds in the cystine knot domain of human von Willebrand factor, *Journal of Biological Chemistry 275*, 25585-25594.

21. Haberichter, S. L., Jacobi, P., and Montgomery, R. R. (2003) Critical independent regions in the VWF propeptide and mature VWF that enable normal VWF storage, *Blood 101*, 1384-1391.

22. Savage, B., Saldivar, E., and Ruggeri, Z. M. (1996) Initiation of platelet adhesion by arrest onto fibrinogen or translocation on von Willebrand factor, *Cell 84*, 289-297.

23. Nishio, K., Anderson, P. J., Zheng, X. L., and Sadler, J. E. (2004) Binding of platelet glycoprotein Ib alpha to von Willebrand factor domain A1 stimulates the cleavage of the adjacent domain A2 by ADAMTS13, *Proceedings of the National Academy of Sciences of the United States of America 101*, 10578-10583.

24. Savage, B., Sixma, J. J., and Ruggeri, Z. M. (2002) Functional self-association of von Willebrand factor during platelet adhesion under flow, *Proceedings of the National Academy of Sciences of the United States of America 99*, 425-430.

25. Shankaran, H., Alexandridis, P., and Neelamegham, S. (2003) Aspects of hydrodynamic shear regulating shear-induced platelet activation and self-association of von Willebrand factor in suspension, *Blood 101*, 2637-2645.

26. Ulrichts, H., Udvardy, M. S., Lenting, P. J., Pareyn, I., Vandeputte, N., Vanhoorelbeke, K., and Deckmyn, H. (2006) Shielding of the A1 domain by the D ' D3 domains of von Willebrand factor modulates its interaction with platelet glycoprotein Ib-IX-V, *Journal of Biological Chemistry 281*, 4699-4707.

27. Choi, H., Aboulfatova, K., Pownall, H. J., Cook, R., and Dong, J. F. (2007) Shear-induced disulfide bond formation regulates adhesion activity of von Willebrand factor, *Journal of Biological Chemistry 282*, 35604-35611.

28. Pimanda, J., and Hogg, P. (2002) Control of von Willebrand factor multimer size and implications for disease, *Blood Reviews 16*, 185-192.

29. Bonnefoy, A., Daenens, K., Feys, H. B., De Vos, R., Vandervoort, P., Vermylen, J., Lawler, J., and Hoylaerts, M. F. (2006) Thrombospondin-1 controls vascular platelet recruitment and thrombus adherence in mice by protecting (sub)endothelial VWF from cleavage by ADAMTS13, *Blood 107*, 955-964.

30. Xie, L. J., Chesterman, C. N., and Hogg, P. J. (2001) Control of von Willebrand factor multimer size by thrombospondin-1, *Journal of Experimental Medicine 193*, 1341-1349.

31. Chopek, M. W., Girma, J. P., Fujikawa, K., Davie, E. W., and Titani, K. (1986) HUMAN VONWILLEBRAND-FACTOR - A MULTIVALENT PROTEIN COMPOSED OF IDENTICAL SUBUNITS, *Biochemistry 25*, 3146-3155.

32. Dong, J. F., Berndt, M. C., Schade, A., McIntire, L. V., Andrews, R. K., and Lopez, J. A. (2001) Ristocetin-dependent, but not botrocetin-dependent, binding

of von Willebrand factor to the platelet glycoprotein Ib-IX-V complex correlates with shear-dependent interactions, *Blood 97*, 162-168.

33. Kasirer-Friede, A., Cozzi, M. R., Mazzucato, M., De Marco, L., Ruggeri, Z. M., and Shattil, S. J. (2004) Signaling through GP Ib-IX-V activates alpha IIb beta 3 independently of other receptors, *Blood 103*, 3403-3411.

34. Christophe, O., Obert, B., Meyer, D., and Girma, J. P. (1991) THE BINDING DOMAIN OF VONWILLEBRAND-FACTOR TO SULFATIDES IS DISTINCT FROM THOSE INTERACTING WITH GLYCOPROTEIN-IB, HEPARIN, AND COLLAGEN AND RESIDES BETWEEN AMINO-ACID-RESIDUES LEU 512 AND LYS 673, *Blood 78*, 2310-2317.

35. Furlan, M., Robles, R., and Lammle, B. (1996) Partial purification and characterization of a protease from human plasma cleaving von Willebrand factor to fragments produced by in vivo proteolysis, *Blood 87*, 4223-4234.

36. Lankhof, H., Damas, C., Schiphorst, M. E., Ijsseldijk, M. J. W., Bracke, M., Furlan, M., Tsai, H. M., deGroot, P. G., Sixma, J. J., and Vink, T. (1997) von Willebrand factor without the A2 domain is resistant to proteolysis, *Thrombosis and Haemostasis 77*, 1008-1013.

37. Lancellotti, S., De Filippis, V., Pozzi, N., Peyvandi, F., Palla, R., Rocca, B., Rutella, S., Pitocco, D., Mannucci, P. M., and De Cristofaro, R. (2010) Formation of methionine sulfoxide by peroxynitrite at position 1606 of von Willebrand factor inhibits its cleavage by ADAMTS-13: A new prothrombotic mechanism in diseases associated with oxidative stress, *Free Radical Biology & Medicine 48*, 446-456.

38. Chen, J. M., Fu, X. Y., Wang, Y., Ling, M. H., McMullen, B., Kulman, J., Chung, D. W., and Lopez, J. A. (2010) Oxidative modification of von Willebrand factor by neutrophil oxidants inhibits its cleavage by ADAMTS13, *Blood 115*, 706-712.

39. Singh, I., Themistou, E., Porcar, L., and Neelamegham, S. (2009) Fluid Shear Induces Conformation Change in Human Blood Protein von Willebrand Factor in Solution, *Biophysical Journal 96*, 2313-2320.

40. Tsai, H. M., Sussman, II, and Nagel, R. L. (1994) SHEAR-STRESS ENHANCES THE PROTEOLYSIS OF VON-WILLEBRAND-FACTOR IN NORMAL PLASMA, *Blood 83*, 2171-2179.

41. Singh, I., Shankaran, H., Beauharnois, M. E., Xiao, Z. H., Alexandridis, P., and Neelamegham, S. (2006) Solution structure of human von Willebrand factor studied using small angle neutron scattering, *Journal of Biological Chemistry 281*, 38266-38275.

42. Kokame, K., Nobe, Y., Kokubo, Y., Okayama, A., and Miyata, T. (2005) FRETS-VWF73, a first fluorogenic substrate for ADAMTS13 assay, *British Journal of Haematology 129*, 93-100.

43. Ruoslahti, E., and Pierschbacher, M. D. (1987) NEW PERSPECTIVES IN CELL-ADHESION - RGD AND INTEGRINS, *Science 238*, 491-497.

44. Voorberg, J., Fontijn, R., Calafat, J., Janssen, H., Vanmourik, J. A., and Pannekoek, H. (1991) ASSEMBLY AND ROUTING OF VONWILLEBRAND-FACTOR VARIANTS - THE REQUIREMENTS FOR DISULFIDE-LINKED DIMERIZATION RESIDE WITHIN THE CARBOXY-TERMINAL-151 AMINO-ACIDS, *Journal of Cell Biology 113*, 195-205.

45. Mayadas, T. N., and Wagner, D. D. (1992) VICINAL CYSTEINES IN THE PROSEQUENCE PLAY A ROLE IN VONWILLEBRAND-FACTOR MULTIMER ASSEMBLY, *Proceedings of the National Academy of Sciences of the United States of America 89*, 3531-3535.

46. Slayter, H., Loscalzo, J., Bockenstedt, P., and Handin, R. I. (1985) NATIVE CONFORMATION OF HUMAN VONWILLEBRAND PROTEIN - ANALYSIS BY ELECTRON-MICROSCOPY AND QUASI-ELASTIC LIGHT-SCATTERING, *Journal of Biological Chemistry 260*, 8559-8563.

47. Emsley, J., Cruz, M., Handin, R., and Liddington, R. (1998) Crystal structure of the von Willebrand factor A1 domain and implications for the binding of platelet glycoprotein Ib, *Journal of Biological Chemistry 273*, 10396-10401.

48. Celikel, R., Ruggeri, Z. M., and Varughese, K. I. (2000) von Willebrand factor conformation and adhesive function is modulated by an internalized water molecule, *Nature Structural Biology 7*, 881-884.

49. Fukuda, K., Doggett, T. A., Bankston, L. A., Cruz, M. A., Diacovo, T. G., and Liddington, R. C. (2002) Structural basis of von Willebrand factor activation by the snake toxin botrocetin, *Structure 10*, 943-950.

50. Huizinga, E. G., Tsuji, S., Romijn, R. A. P., Schiphorst, M. E., de Groot, P. G., Sixma, J. J., and Gros, P. (2002) Structures of glycoprotein Ib alpha and its complex with von Willebrand factor A1 domain, *Science 297*, 1176-1179.

51. Huizinga, E. G., Martijn van der Plas, R., Kroon, J., Sixma, J.J., Gros, P. (1997) Crystal structure of the A3 domain of human von Willebrand factor: implications for collagen binding., *Structure 5*, 1147-1156.

52. Zhang, Q., Zhou, Y. F., Zhang, C. Z., Zhang, X. H., Lu, C. F., and Springer, T. A. (2009) Structural specializations of A2, a force-sensing domain in the ultralarge

vascular protein von Willebrand factor, *Proceedings of the National Academy of Sciences of the United States of America 106*, 9226-9231.

53. Luken, B. M., Winn, L. Y. N., Emsley, J., Lane, D. A., and Crawley, J. T. B. (2010) The importance of vicinal cysteines, C1669 and C1670, for von Willebrand factor A2 domain function, *Blood 115*, 4910-4913.

54. Doggett, T. A., Girdhar, G., Lawshe, A., Schmidtke, D. W., Laurenzi, I. J., Diamond, S. L., and Diacovo, T. G. (2002) Selectin-like kinetics and biomechanics promote rapid platelet adhesion in flow: The GPIb alpha-vWF tether bond, *Biophysical Journal 83*, 194-205.

55. Yago, T., Lou, J., Wu, T., Yang, J., Miner, J. J., Coburn, L., Lopez, J. A., Cruz, M. A., Dong, J. F., McIntire, L. V., McEver, R. P., and Zhu, C. (2008) Platelet glycoprotein lb alpha forms catch bonds with human WT vWF but not with type 2B von Willebrand disease vWF, *Journal of Clinical Investigation 118*, 3195-3207.

56. Siedlecki, C. A., Lestini, B. J., KottkeMarchant, K., Eppell, S. J., Wilson, D. L., and Marchant, R. E. (1996) Shear-dependent changes in the three-dimensional structure of human von Willebrand factor, *Blood 88*, 2939-2950.

57. Auton, M., Zhu, C., and Cruz, M. A. (2010) The Mechanism of VWF-Mediated Platelet GPIb alpha Binding, *Biophysical Journal 99*, 1192-1201.

CHAPTER 2
EXPRESSION AND PURIFICATION OF RECOMBINANT VON WILLEBRAND FACTOR A1A2A3 DOMAINS

ABSTRACT

In order to initiate the formation of a platelet plug VWF must be assembled into large multimers. VWF undergoes post translational modifications by dimerizing through multiple intermolecular disulfide bonds between carboxyl terminal ends of the protein and once in Golgi by forming interdimer disulfide bonds. The resulting multimers range in size between 500 to 20000 kDa. The protein dimerizes and the dimers then form a variety of disulfide crosslinked multimers with as many as 80 monomeric units, weighing more than 20 million Daltons. Studying such an enormous molecule poses special challenges. The separate domains within the VWF subunit exhibit specific properties, involving interactions with other molecules. Binding sites that are independent of multimer assembly but important for the hemostatic function are located in the A1A2A3 domains of VWF. We expressed the A1, A2, and A3 domains of von Willebrand factor in a single polypeptide using *Pichia Pastoris* expression system. Proteins with disulfide bonds, requiring post translational modifications and glycosylation can be produced in their correctly native folded states with full function from *Pichia pastoris*. We purified the A1A2A3 domain using ethanol, ammonium sulfate precipitation and ion exchange chromatography. Our efforts in solubilizing the purified protein were unsuccessful more likely due to the unusual adhesive nature of the A1A2A3 domain of the VWF.

INTRODUCTION

Von Willebrand factor (VWF) is a large glycoprotein, found in plasma and platelets and synthesized by megakaryocytes and endothelial cells. The pre-pro-protein is 2813 amino acids and the mature monomer is 2050 amino acids. VWF monomer contains multiple copies of four types of domains called A, B, C and D-type domains. Binding sites that are independent of multimer assembly but important for the hemostatic function are located in the A1A2A3 domains of VWF.

Expression of the A1, A2, and A3 domains individually in *E. coli* has been successful, but attempts to express all three in a single polypeptide in E. coli have failed. We attempted to express, purify, and characterize the A1, A2, and A3 domains of von Willebrand factor in a single polypeptide in order to study a model more realistic than a simple peptide. The lack of domains responsible for the multimer formation would make it more practical to study and less complex than multimers ranging in size from 2-20 million Daltons. Proteins with post translational modifications with disulfide bonds and requiring glycosylation are very difficult to obtain using E. coli expression, *Pichia pastoris* has been used successfully to produce a number of difficult to express proteins in large amounts (*1*). Formerly in our lab the portion of the human gene expressing the A1, A2, and A3 domains was subcloned into an E. coli vector and homologously recombined into a site in *Pichia pastoris.* Unfortunately, this failed to express, then a synthetic gene was optimized in codon usage and homologously recombined into *Pichia.*

Figure 2.1. Amino acid sequence of VWF A1A2A3 used in Pichia pastoris expression studies.

Translate from the Codon Optimized Synthetic Gene

EFDISEPPLHDFYCSRLLDLVFLLDGSSRLSEAEFEVLKAFVVDMMERLRISQKWVRVAV
VEYHDGSHAYIGLKDRKRPSELRRIASQVKYAGSQVASTSEVLKYTLFQIFSKIDRPEAS
RIALLLMASQEPQRMSRNFVRYVQGLKKKKVIVIPVGIGPHANLKQIRLIEKQAPENKAF
VLSSVDELEQQRDEIVSYLCDLAPEAPPPTLPPHMAQVTVGPGLLGVSTLGPKRNSMVLD
VAFVLEGSDKIGEADFNRSKEFMEEVIQRMDVGQDSIHVTVLQYSYMVTVEYPFSEAQSK
GDILQRVREIRYQGGNRTNTGLALRYLSDHSFLVSQGDREQAPNLVYMVTGNPASDEIKR
LPGDIQVVPIGVGPNANVQELERIGWPNAPILIQDFETLPREAPDLVLQRCCSGEGLQIP
TLSPAPDCSQPLDVILLLDGSSSFPASYFDEMKSFAKAFISKANIGPRLTQVSVLQYGSI
TTIDVPWNVVPEKAHLLSLVDVMQREGGPSQIGDALGFAVRYLTSEMHGARPGASKAVVI
LVTDVSVDSVDAAADAARSNRVTVFPIGIGDRYDAAQLRILAGPAGDSNVVKLQRIEDLP
TMVTL-AAAVD

EF at N-terminus is due to restriction site/signal peptide.
The third amino acid, D, is residue 1261 in pre-pro-von Willebrand factor.
The L just before the stop codon is residue 1863.
Estimated pI =5.29 at pH 7.0
N indicates N-glycosylation site. YM is ADAMTS-13 cleavage site
Isotopically Averaged Molecular Weight = 66755.3359

We initially focused on the expression and the purification of the recombinant A1A2A3 domain of von Willebrand factor . Our plan was to characterize the redox dependence of the disulfide bonds in the A1A2A3 domain of von Willebrand factor under oxidized and reduced glutathione buffer systems.

MATERIALS AND METHODS

Growth and maintenance of P. pastoris strains

The codon optimized (Figure 2.1.) A1A2A3 domain of VWF was expressed in the *Pichia pastoris* expression system. The protease A, carboxypeptidase Y and protease B1 deficient SMD1168 strain of *P. pastoris* was used for the production of the recombinant protein. The use of protease-deficient strains such as SMD1168 has been found to improve both yield and quality of the recombinant protein. The SMD1168 strain of *Pichia pastoris* is defective in the histidine dehydrogenase gene (his4), which allowed the selection of transformants based of their ability to grow on histidine-deficient medium. *P. pastoris* strains were typically grown on YPD-agar plates (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose and 2% (w/v) agar) or YPD medium (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose). One liter solution of YPD media was prepared by dissolving all the ingredients in 900ml of reverse osmosis water and sterilizing by autoclave process, the dextrose was dissolved in 100ml of reverse osmosis water and autoclaved separately. After cooling the YPD solution to 55 ºC, 100 ml dextrose solution ascetically transferred and the medium plated in sterile Petri plates. To prepare cells for long term storage, a single colony selected from YPD-agar plates was used to inoculate a 10 ml solution of YPD in a sterile baffled flask. The culture allowed to shake (250 rpm) at 30 ºC for 24h. The cells were collected by centrifugation (Sorvall, GS-3 rotor) for 10 min at 3000 rpm and resuspended in a 1ml solution of YPD containing 15% glycerol. The cultures were then flash-frozen in liquid nitrogen and stored at -80 ºC. These ‘frozen stocks’ were used to inoculate fresh cultures(*2*).

Protein Production

Cultures of *Pichia pastoris* were grown in BMGY, composed of 3.4g/L yeast nitrogen base without amino acids and ammonium sulfate, 1% (w/v) ammonium sulfate, 1% (w/v) casamino acids, 2ml/L biotin solution, 2% (v/v) 1M potassium phosphate buffer (pH 6.0, adjusted with potassium hydroxide) and 1% (v/v) glycerol. The prepared BMGY solution sterilized by filtration through a 0.22µm PES filter (Corning Incorporated). The biotin solution was prepared by dissolving 0.2g/L of biotin in 0.02M potassium hydroxide solution, then sterilized by filtration through a 0.22µm PES filter (Corning Incorporated) and stored in aliquots at -20 ºC. Prior to inoculation, 2L culture flasks covered with four layers of cheese cloth and two layers of aluminum foil then sterilized by autoclaving at 121 ºC for 20 min. The sterile BMGY solution was transferred aseptically into the 2L culture flasks. The amount of media in the flasks was usually $1/5^{th}$ of the flask volume, 4L of BMGY was equally distributed into 12 sterile 2L culture flasks. A starter culture of *P. pastoris* was prepared by inoculating 10ml of BMGY in a sterile baffled flask with the frozen stock culture and allowed to incubate at 30 ºC with shaking (250 rpm) for 12h. The media in the 2L culture flasks was inoculated with the 10 ml starter culture. After 24h of growth, the aluminum foil on the 2L culture flasks was removed. After this point forward, the flasks were covered with only cheese cloth. This allows better aeration and contamination is not a concern in actively growing yeast cultures. After 48h of total growth in the 2L culture flasks, the cells were harvested by centrifugation at 3000 rpm for 10 minutes. The supernatant was discarded and the cells were resuspended in BMMY. Since mut^S phenotype of *Pichia* was utilized, the volume of BMMY used was half the volume of BMGY utilized during the growth phase.

BMMY has the same composition as BMGY except 2% methanol was used as a replacement for 1% glycerol. The cells resuspended in BMMY were returned to the 2L culture flasks and covered with cheese cloth and incubated with shaking (250 rpm) at 30 ºC. Induction periods was 36h. Cells were harvested by centrifugation at 3000 rpm for 60 min. The cells were discarded and 5mM disodium EDTA to final concentration was added to the supernatant. The supernatant containing the secreted protein may be frozen at – 20 ºC for several weeks before purification.

Protein Purification

The proteins are least soluble when at their isoelectric point thus the pH of supernatant adjusted to 5.2 using concentrated glacial acetic acid. The pI of A1A2A3 domains of VWF was calculated to be 5.4. Solid ammonium sulfate was added to supernatant to a final molar concentration of 2.5M, initially. The supernatant was centrifuged (Sorvall, GS-3 rotor) at 5000 rpm for 10 min. Pellets corresponding to 2.5M ammonium sulfate precipitation was discarded then solid ammonium sulfate is added to supernatant to a final molar concentration 5M. The solution was allowed to equilibrate for 1h at 4 ºC. The supernatant was centrifuged (Sorvall, GS-3 rotor) at 5000 rpm for 10 min, pellet resuspended in 50mM MES, 1mM EDTA at pH 6.5. Then the final volume of cold ethanol adjusted to 80% and incubated at -20 ºC for 1 hour, then centrifuged (Sorvall, GS-3 rotor) at 5000 rpm for 10 min. The supernatant was discarded. The pellet resuspended in 50mM MES, 1mM EDTA at pH 6.5. Approximately 2.0 ml of the 80% ethanol precipitation pellet was dialyzed overnight at 4 ºC against 1L. 25mM MES, 1mM EDTA at pH 6.5, dialysis buffer was changed twice. Two milliliters of Sepharose fastflow Q (GE Healthcare) was poured into a 12 ml Bio-Rad disposable column. The

column was equilibriated with 10ml of 25mM MES, 1mM EDTA at pH 6.5. Dialyzed 2 ml sample then passed over the bed followed by washing with an additional 2ml 25mM MES, 1mM EDTA at pH 6.5. To elute the protein the column was washed with 4ml of 1M NaCl, 25mM MES, 1mM EDTA at pH 6.5. The elution product was concentrated down to 1-2 ml volume using Amicon Ultra–15, regenerated cellulose (30000 kDa molecular weight cutoff) from Millipore corporation The sample was stored at -20 ºC or immediately subject to HPLC. The concentrated sample then injected into a Waters HPLC system for reverse phase chromatography containing a Atlantis dC18, 5µm particle, 300 A° 4.6x250mm column held at 58°C. Upon injection a gradient is run at a flow rate 1 ml/min from 10 to 90% acetonitrile, 01% trifluroacetic acid over 37 minutes.

Western Blot Analysis

Protein samples were resolved by 12.5% SDS–PAGE under reducing conditions using duplicate gels. One gel was Coomassie Blue stained, the other was used for electrophoretic transfer (Trans-Blot SD Semi-Dry Cell apparatus, Bio-Rad) of the eluted and or fractionated proteins onto PVDF membranes. The membranes were incubated for one hour at room temperature in 3% blocking reagent to prevent non- specific binding. Then membrane was incubated overnight at 4 ºC with 1:1000 dilution of primary antibody (Rabbit polyclonal to von Willebrand Factor), then 1:2000 dilution of HRP conjugated secondary antibody (Donkey polyclonal to Rabbit IgG) . Membranes were washed in PBST for one hour at room temperature and signals developed following the (Bio-Rad) chemiluminescent kit protocol.

Deglycosylation

For mass spectrometry analysis, the protein was treated with Deglycosylation Enzyme Mix (Endo-α-N-Acetylgalactosaminidase, PNGase F,Neuraminidase, β1-4 Galactosidase, β-N-Acetylglucosaminidase) (New England Biolabs, Ipswich,MA). Protein deglycosylation enzyme mix was supplied in 50 mM NaCl and 20 mM Tris-HC l at pH 7.5. Protein deglycosylated by the addition of 10 Units of enzyme mix followed by incubation at 37°C for four hours.

Mass Spectrometry Analysis

The identity of the protein was confirmed using MALDI-TOF. Samples were analyzed at the Statewide Mass Spectrometry Laboratory located at the University of Arkansas. Two microliter of deglycosylated protein sample was mixed with an equal volume of 1M 2, 5-dihydroxybenzoic acid (DHB) in 90% methanol containing 0.1% formic acid and spotted onto a Bruker MTP 384 stainless steel MALDI target. MALDI-TOF analysis was performed using a Bruker Reflex III MALDI-TOF mass spectrometer (Bruker Daltonik GMBH, Bremen, Germany). Mass spectrometer operated in positive ion reflectron mode.

Digestion of Recombinant VWF A1A2A3

Protein samples were deglycosylated using Deglycosylation Enzyme Mix (New England Biolabs) as described previously. Sequencing grade Trypsin or Pepsin (Promega, WI) prepared according to manufacturer's instructions. Protein was taken into the buffers listed in the second column of Table 2.1. Buffer exchange was performed according to the manufacturer's instructions using Amicon Ultra–15, regenerated cellulose (10000 kDa molecular weight cutoff) centrifugal filters. Reduction and Alkylation carried as part of our efforts to digest the protein. Protein sample in 25 mM

ammonium bicarbonate subjected to reduction and alkylation prior to digestion with recombinant trypsin. Sample reduced for 1 hr with 10 mM 1,4- dithiothreitol (DTT) followed by alkylation with 20 mM 2-iodoacetamide in dark for 20 min at 37°C. Excess iodoacetamide was quenched by an addition of DTT.

Reduced and alkylated peptides were then digested with trypsin for 24 h at 37°C, desalted, and concentrated using C18 ZipTips (Millipore) as recommended by manufacturer's protocol and spotted on to the MALDI target plate for analysis.

Table 2.1. Digestion Conditions

Sample	Buffer	Reduction/Alkylation	Denaturant/Surfactant	Enzyme
10µl.(0.9mg/ml)	PBS	TCEP	-	Trypsin
HPLC fraction	25mM NH_4HCO_3	-	-	Trypsin
HPLC fraction	25mM NH_4HCO_3	10 mM DTT, 20 mM Iodoacetamide	-	Trypsin
In-Gel Digestion	25mM NH_4HCO_3	10 mM DTT, 20 mM Iodoacetamide	-	Trypsin
10µl.(0.9mg/ml)	25mM MES,1M NaCl	-	-	Trypsin
10µl.(0.9mg/ml)	25mM NH_4HCO_3	10 mM DTT, 20 mM Iodoacetamide	-	Trypsin
10µl.(0.9mg/ml)	25mM NH_4HCO_3	10 mM DTT, 20 mM Iodoacetamide	6M Urea	Trypsin
10µl.(0.9mg/ml)			60% Acetonitrile	Trypsin
10µl.(0.9mg/ml)	PBS	-	0.1%Tween 20	Trypsin
10µl.(0.9mg/ml)	Sodium Acetate	-	-	Pepsin
10µl.(0.9mg/ml)	Sodium Acetate	10 mM DTT, 20 mM	-	Pepsin

RESULTS AND DISCUSSION

The goal of this study was to purify 5-10 mg of recombinant VWF A1A2A3 domain using *Pichia pastoris* expression system. Expression of recombinant A1A2A3 in *Pichia* was determined using Western Blot assays (Figure 2.2)

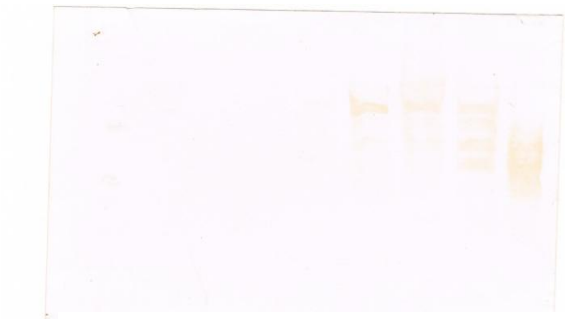


Figure 2.2 Western Blot from the crude supernatant.

Following chromatogram (Figure 2.3) was obtained when the ammonium sulfate and ethanol precipitation product was eluted using 1M NaCl 50 mM MES 10mM EDTA, pH 6.5, 100µl was injected in a 10 to 90 % acetonitrile gradient.

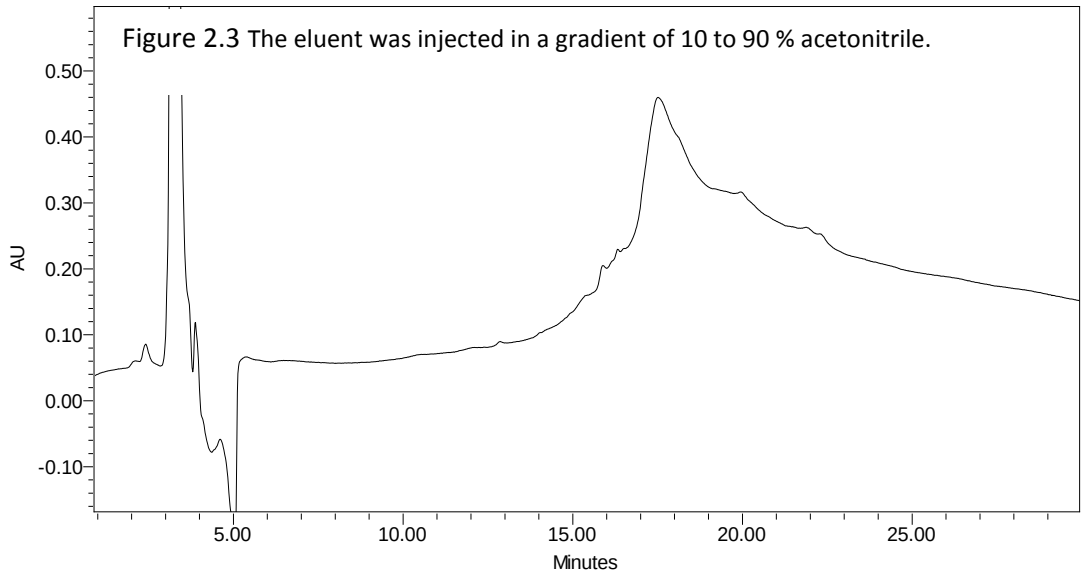


Figure 2.3 The eluent was injected in a gradient of 10 to 90 % acetonitrile.

1ml/min fractions were collected flash-frozen in liquid nitrogen and lyophilized using SpeedVac, reconstituted with 50 µl with deonized water and sonicated. Fractions corresponding to 15 to 21 minutes were submitted for mass spectrometry analysis.

The following mass spectrum (Figure 2.4) was obtained from the HPLC fraction corresponding to 19th minute by MALDI TOF. This protein has a molecular weight near the outer limit of the capability of this particular instrument. There is also a possibility that the measured molecular weight is higher because of tight binding metallic cations. With these caveats, Dr. Lay estimates the molecular weight to be 66.1 ±0.3 kDa. The predicted molecular weight for the expected polypeptide is 66,701.

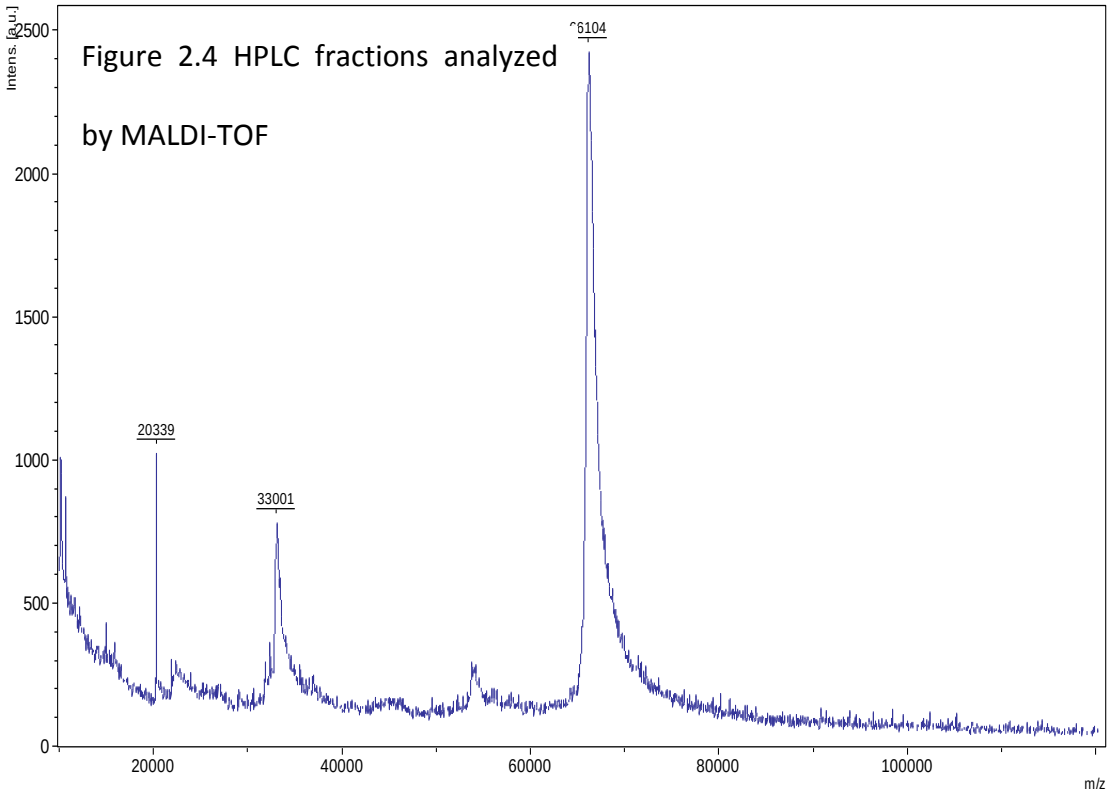


Figure 2.4 HPLC fractions analyzed by MALDI-TOF

When we examined 35µl of the sample on a SDS gel a single band corresponding to approximately 65kDa was observed. (Figure 2.5).

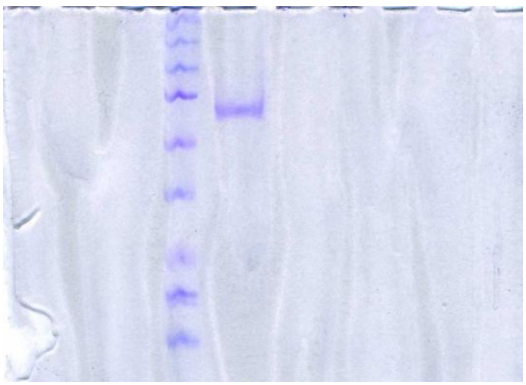


Figure 2.5 12.5% SDS gel for the final purification of recombinant VWF A1A2A3 domains

Same sample developed a positive signal on Western blot at approximately at 65 kDa position (Figure 2.6).

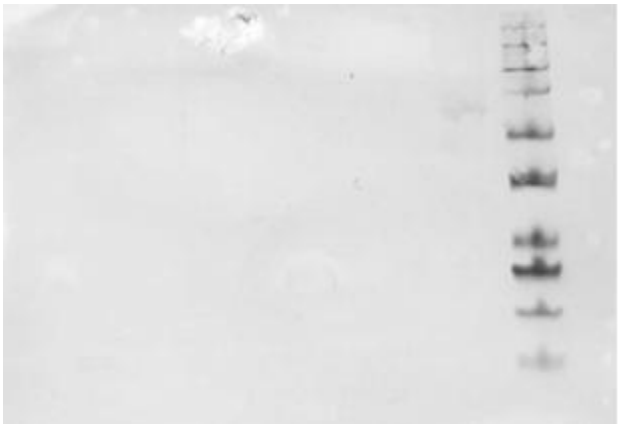
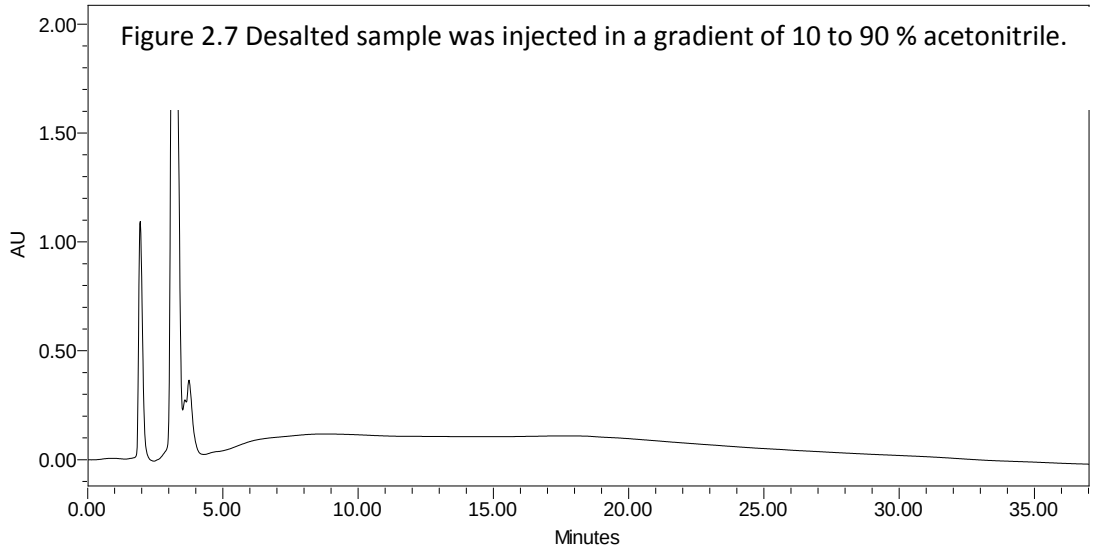
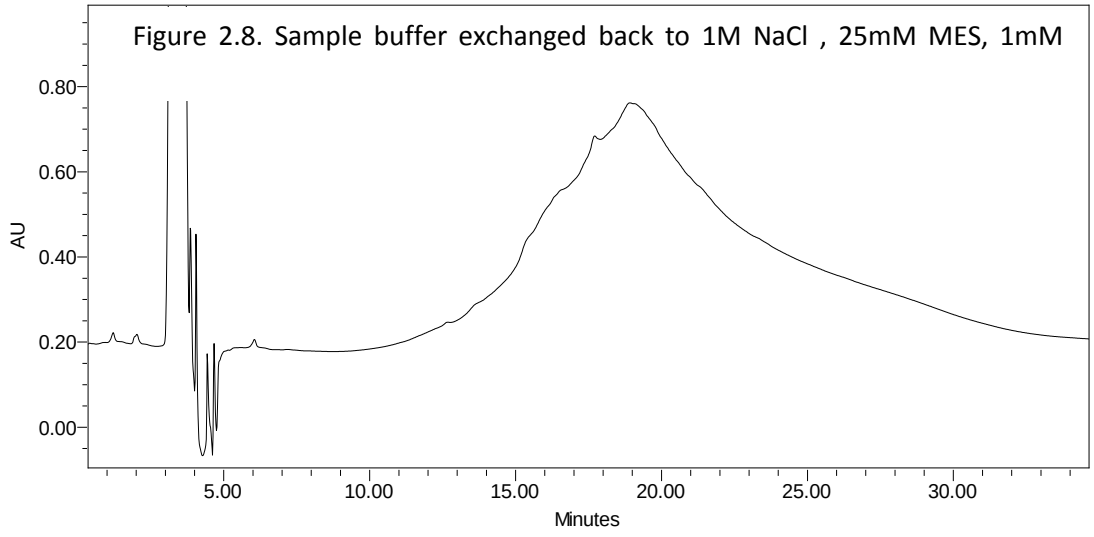


Figure 2.6. Western Blot for the final purification of recombinant VWF A1A2A3 domains

Desalting of protein in 1M NaCl , 25mM MES, 1mM EDTA solution was necessary to conduct the experiment protocols for the characterization of the disulfide bonding pattern of the purified recombinant protein. When we buffer exchanged the sample into 25mM MES, 1mM EDTA with PBS following chromatogram was obtained (Figure 2.7).



When we buffer exchanged back to 1M NaCl , 25mM MES, 1mM EDTA we obtained the following chromatogram (Figure 2.8).



N-terminal sequencing attempts to determine the N-terminal sequence of the protein with Edman degradation have failed. We were unable to digest the protein despite of all efforts summarized in Table.2.1. The A1A2A3 domain of VWF is adhesive in nature, contains binding sites for platelets, heparin and collagen. In our experiments we

successfully expressed and purified the recombinant protein but we were unable to solubilize the protein, in the absence of at least 1M NaCl, in standard buffer solutions such as phosphate buffered saline (PBS; 1.46 mM KH2P04, 9.9 mM Na2HP04, 2.68 mM KCl, 137 mM NaCl, pH 7.4). Protein purifications performed using described protocol, in various fermentation batches typically yielded 0.3 to 0.9 mg of recombinant VWF, when we determined the protein concentration by absorbance at 280 nm. We were unable to solubilize the recombinant protein in conditions that allow for subsequent studies. Solubility became our challenge more likely due to the lack of the domains that contributes to the solubility of the VWF under physiological conditions. In order to overcome this challenge we tried to solubilize the protein in an acidic environment below the pI of the A1A2A3 domain, using Sodium acetate buffer at pH 4. We were unable to digest the protein under these conditions as well both using Trypsin and Pepsin enzymes. We also tried to denature the protein prior to digestion using urea, tween 20 and acetonitrile we were unable digest the protein under these conditions probably due to the interfering effect of the denaturants with the enzymes. In the future efforts residues that could enhance solubility could be included to the synthetic gene of the recombinant protein.

REFERENCES

1. Cregg, J. M., Vedvick, T.S. and Raschke, W.C. . (1993) Recent advances in the expression of foreign genes in Pichia pastoris. , *Biotechnology 11*, 905-910.

2. Talla, D. (2007) Heat capacity changes in staphylococcal nuclease packing mutants and preliminary development of an immunoassay for thrombomodulin. , In *Biochemistry*, University of Arkansas, Fayetteville.

CHAPTER 3
REDOX DEPENDENT DISULFIDE BOND FORMATION IN VON WILLEBRAND FACTOR A1A2A3 DOMAINS

ABSTRACT

Von Willebrand factor (VWF) is a large glycoprotein that plays an important role in primary hemostasis by forming a bridge between subendothelial matrix and platelets. The most striking feature of VWF is its ability to dramatically shift from a non-adhesive plasma soluble protein to a multi-functional adhesive protein, making an activated surface, attractive to the flowing platelets. The segmental structure of VWF forms the basis of its multifunctional properties. Binding sites that are independent of multimer assembly but important for the hemostatic function are located in the A1A2A3 domains of VWF. VWF's A1A2A3 domains contain six cysteine residues in three disulfides, one being a vicinal disulfide. The established view is that disulfide bonds are inert motifs, stabilizing the protein structure. However vicinal disulfides in other proteins have been reported to function as redox activated conformational switches. In order to explore the possibility of redox dependent disulfide bonding of the VWF A1A2A3 domains, we re-examined disulfide bonding in both reduced and oxidized glutathione environments. Disulfide mapping was performed by digesting the A1A2A3 domain using endoprotease GluC and analyzing the digestion products by using matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS). The best interpretation of the MALDI-MS data is that disulfide bonds different from those previously reported in the literature are present in VWF. In particular a disulfide bond can be formed between the A1 and the A2 domains.

INTRODUCTION

Disulfide bond formation is a reversible reaction in which the thiol groups of two cysteine residues are oxidized to form a covalently linked disulfide.

$$\text{R-Cys-SH + R'-Cys-SH} \leftrightarrows \text{R-Cys-S-S-Cys-R' + 2 } e^- + 2\ H^+$$

This linkage can be constructed by thiol-disulfide exchange in which a disulfide bond, present in an electron acceptor, is reduced (*1*) . The formation of a disulfide bond can enhance stability and facilitate protein folding. Disulfide bonds stabilize protein folding via several mechanisms. They firmly hold two or more distal portions together and contribute to the stabilization of folded topology. Disulfide bonds link two or more regions of the protein chains, and thus increase the effective local concentration of protein residues and lower the local concentration of water molecules. Disulfide bonds also can serve as the center of hydrophobic cores of the folded proteins. Disulfide bonds stabilize alpha helical and beta structures by preventing access of water to attack the hydrogen bonds of residues (*1*, *2*).

Disulfides can affect protein folding in another way. The Anfisen experiment proved that the three-dimensional structure and the function of a protein are determined by its amino acid sequence (*3*). The native structure in general represents the lowest free energy state of a polypeptide chain. However, oxidative refolding of reduced protein *in vitro* is observed to be slow and error prone, resulting in formation of incorrect disulfide bonds which are slow to convert to other forms. In order to prevent this low molecular weight thiols such as glutathione can be added in both their reduced and oxidized form as

"redox shuffling" reagents. "Redox-shuffling" reagents enable thiol-disulfide exchange reactions in both directions, thus reshuffling improper disulfide bonds and increasing the yield of proteins with the correct disulfide bonds(*4, 5*).

The established view has been that disulfide bonds are inert motifs which once formed would not be altered easily, but this opinion is changing rapidly(*6, 7*). It has been found that in some proteins thiols and disulfide bonds can undergo redox state changes under physiological conditions (*6, 8*) that are important to their function. For example, it has been proposed that chemical reactions involving Cys186 and Cys209 in the extracellular domain of tissue factor are the key events regulating TF procoagulation function. When the side chains of cysteines are reduced TF is inactive. When the sulfhydryls are oxidized the resulting disulfide bond formation induces a conformational change that converts TF to a procoagulant state (*9-12*). In another study Stathakis et al. (*13*) showed that the disulfide bond Cys 461-Cys540 in the kringle V domain of plasminogen is reduced during the process of angiostatin production. According to the authors, thrombus formation is likely redox regulated and involves disulfide bond reduction. In several systems, disulfide bonds between vicinal cysteines have been reported to cycle between reduced and oxidized states (*14*). The crystal structure of VWF A2 domain revealed the presence of two vicinal Cys1669-Cys1670 residues that are linked by a disulfide bond. The peptide backbone of Cys1669–Cys1670 and the vicinal disulfide forms an eight-membered ring (*15*). It has been proposed that this ring motif in vicinal disulfide bonds contributes to a possible conformational switch function, controlling active and inactive states of the protein (*16*). A regulatory role for this motif

has been reported for enzymes such as methanol dehydrogenase, mercuric ion reductase and transglutaminase (*17-19*).

It has been recently reported that disulfide interchanges also occur in VWF (*20*). According to Choi et al. (*21*) shear rates in range of 2000-5000/s causes novel disulphide bond formation in VWF and shear-induced thiol-disulfide exchange increases VWF binding to platelets. The thiol disulfide exchange specifically involves the cysteine residues in the D3 and C domains as determined by mass spectrometry of the tryptic VWF peptides. ADAMTS-13 has reported to have a disulfide bond reducing activity that regulates the shear-induced thiol-disulfide exchange and the disulfide-bond-reducing activity of ADAMTS-13 may prevent platelet adherence of VWF multimers induced by high fluid shear stress by targeting the Cysteine thiols that are in the VWF C-domain (*22*). The research conducted on ADAMTS-13 knock-out mice and patients with near complete absence of ADAMTS-13 activity revealed that ultra large VWF does not build up to levels necessary to trigger disease, and other regulators of VWF multimer size may be involved. Xie et al. (*23*) reported that thiol-disulfide exchange could be relevant for the main control of VWF multimer size, while ADAMTS-13 cleavage may be more functional under pathogenic conditions for endothelial/ platelet activation. According to Xie et al. (*24*), VWF multimer size is under redox regulation. Reduction in VWF multimer size is known to be associated with formation of thiol-dependent complexes of thrombospondin. These observations associated thrombospondin-1 with a disulfide bond reductase- isomerase activity. Reductase-isomerase active sites are characterized by a motif of cysteine residues in the consensus sequence; CGXC. The cysteine thiols cycle

between the reduced and oxidized disulfide bond in coordination with a dithiol or disulfide of a protein substrate which can result in formation of disulfide bonds in the protein substrate (*25*).

The plasma redox potential may change in response to several pathological conditions any consequent changes in VWF would have the potential to alter the hemostatic balance. It remains to be resolved whether redox regulation of VWF in plasma can occur. Other than proteins such as disulfide isomerases, small redox molecules such as GSH-GSSG might also be involved in controlling VWF multimer size (*7*).

Redox control of protein structure and function.

Thermodynamic stability of disulfide bonds can be measured in terms of their redox potential, their potential for receiving electrons. A lower redox potential corresponds to more stable disulfide bonds. The redox potential, E_h is a quantitative expression of the electromotive force (reducing force) of a reductant-oxidant pair expressed relative to a standard hydrogen electrode, providing information about the thiol/disulfide interactions within the environment they create. The redox potential involves the concentrations of the thiols and disulfides and their ability to donate and accept electrons, and the pH of the environment. The Nernst equation is the quantitative expression of the redox potential.

$$E_h = E_0 + \frac{RT}{nF} ln \frac{electron\ acceptor}{electron\ donor}$$

In this equation R is the gas constant, T is the temperature, F is Faraday constant, E_0 is the standard potential relative to a standard hydrogen electrode. The logarithmic term expresses the dependence of the redox potential on concentrations of the oxidized and reduced forms.

Two systems maintain the cellular redox homeostasis are the thioredoxin system and the glutaredoxin system. Both systems use NADPH as the primary hydride donor. The components of the glutaredoxin system are glutaredoxin, glutathione, NADPH, and glutathione reductase. In the glutaredoxin system electrons are transferred from NADPH, through glutathione reductase, which reduces its substrate oxidized GSSG to two molecules of glutathione GSH. The thioredoxin system is also responsible for maintaining the reduced cellular environment. This system is composed of thioredoxin and thioredoxin reductase. Oxidized thioredoxin is reduced to its thiol form by thioredoxin reductase and NADPH (*26*).

Glutathione is a tripetide composed of glutamate, glycine and cysteine. The nucleophilic, and reductant capabilities of GSH are conferred by its sulfhydryl group. Glutathione is the most abundant non-protein thiol found in cells. After donating an electron from its thiol group, GSH itself becomes reactive and readily forms a disulfide bond with another GSH molecule to form GSSG, or oxidized glutathione (*27*).

The GSH-GSSG redox couple is important in regulating cellular proteins through reversible disulfide bond formation. The formation of inter- and intramolecular disulfides between protein SH-groups and GSH is called protein glutathiolation (*28*). Maintaining a normal GSH homeostasis is important for critical cell functions,

imbalances in cellular GSH levels are observed in cancer, neurodegenerative diseases, AIDS, aging autism and autoimmune diseases (*29, 30*).

The major difference between the intracellular and extracellular environments in terms of their relative redox states is that the majority of the low molecular weight thiol/disulfide pairs in cells are GSH/GSSG whereas the majority low molecular weight system in the extracellular environment is Cys/CysSSCys. Although under normal conditions cysteine/cystine concentrations are higher than GSH concentrations in plasma, cysteine is toxic to cell at high concentrations. GSH, serving as a cysteine source, can be degraded to release cysteine when needed. In human blood the majority of glutathione is found inside the erythrocytes. In the erythrocytes, glutathione is the main redox buffer, responsible of the maintenance of iron in its reduced form, a prerequisite for hemoglobin's function (*31-33*).

Glutathione is still found extracellularly and, for example, is present in plasma, the lumen of small intestine, alveolar lining fluid, seminal fluid, and cerebrospinal fluid (*34*). One of the functions of glutathione is the maintenance of extracellular thiol disulfide balance to keep receptors, enzymes, transporters in their functional state. Extracellular GSH is the storage form of cysteine and serves as a source of cysteine (*35*). In the plasma of humans, GSH concentration is in the range of 2-4μM with a redox value of 137± 9 mV (*36-38*). Plasma GSH/GSSG hoemostasis is determined by hepatic output and peripheral uptake. The main source of plasma GSH is hepatic GSH, exported across the hepatic sinusoidal membrane (*35, 36*). Hepatic GSH synthesis is under hormonal control and the availability of methionine and cysteine is also a limiting factor for GSH

synthesis. Insulin and glucagon hormones stimulate the GSH efflux. GSH uptake through intestine also contributes to the plasma GSH pool (*39*).

Control of plasma GSH status is a function of GSH involvement in redox reactions as well as peripheral tissue uptake. GSH functions in plasma to maintain plasma proteins in their functional form. It has been shown that blood clotting is sensitive to the glutathione redox couple *in vitro*. When clotting assays were established under various redox conditions, it has been shown that GSH-GSSG interchange shown to function in the integration and control of the coagulation cascade (*40*) . Plasma glutathione plays a role to eliminate plasma lipid peroxidases, to supply GSH to cells for GSH dependent detoxification, to reduce oxidants, to maintain plasma antioxidant molecules α-tocophenol and ascorbate in their reduced form (*41*). Changes in the plasma GSH-GSSG pool have been shown to be associated with oxidative disturbance in specific organs or the whole body, advanced aging, chronic hypoxia, cardiovascular disease, diabetes, cirrhosis, and HIV infection (*29, 30, 42, 43*).

Disulfides in von Willebrand factor.

A VWF monomer consists of twelve domains; D1, D2, D', D3, A1, A2, A3, D4, B, C1, C2, and CK. The only domain, independent of multimer assembly in VWF is the A domain, which is the focus of this study. Two VWF monomers form a dimer through disulfide bonds between two CK domains. Then dimers are polymerized into high molecular weight multimers, by disulfide bonds connecting their D'D3 domains. Twenty-eight disulfide bonds of VWF have been reported in the literature (Figure 3.1). Inter-subunit disulfide bonds have been reported to be localized in the A and CK domains.

Binding sites that are important for the hemostatic function are located in the A1A2A3 domains of VWF. VWF A1A2A3 domain contains six cysteine residues in three disulfides, one being a vicinal disulfide. Both the A1 and A3 domains form a large loop that is fortified at the base by an intra-chain disulfide bond between cysteines, as shown in Figure 3.2 and Figure 3.3.

Figure 3.1. Disulfide bonds of VWF (Generated with Bruker Daltonics Sequence Editor 3.1)

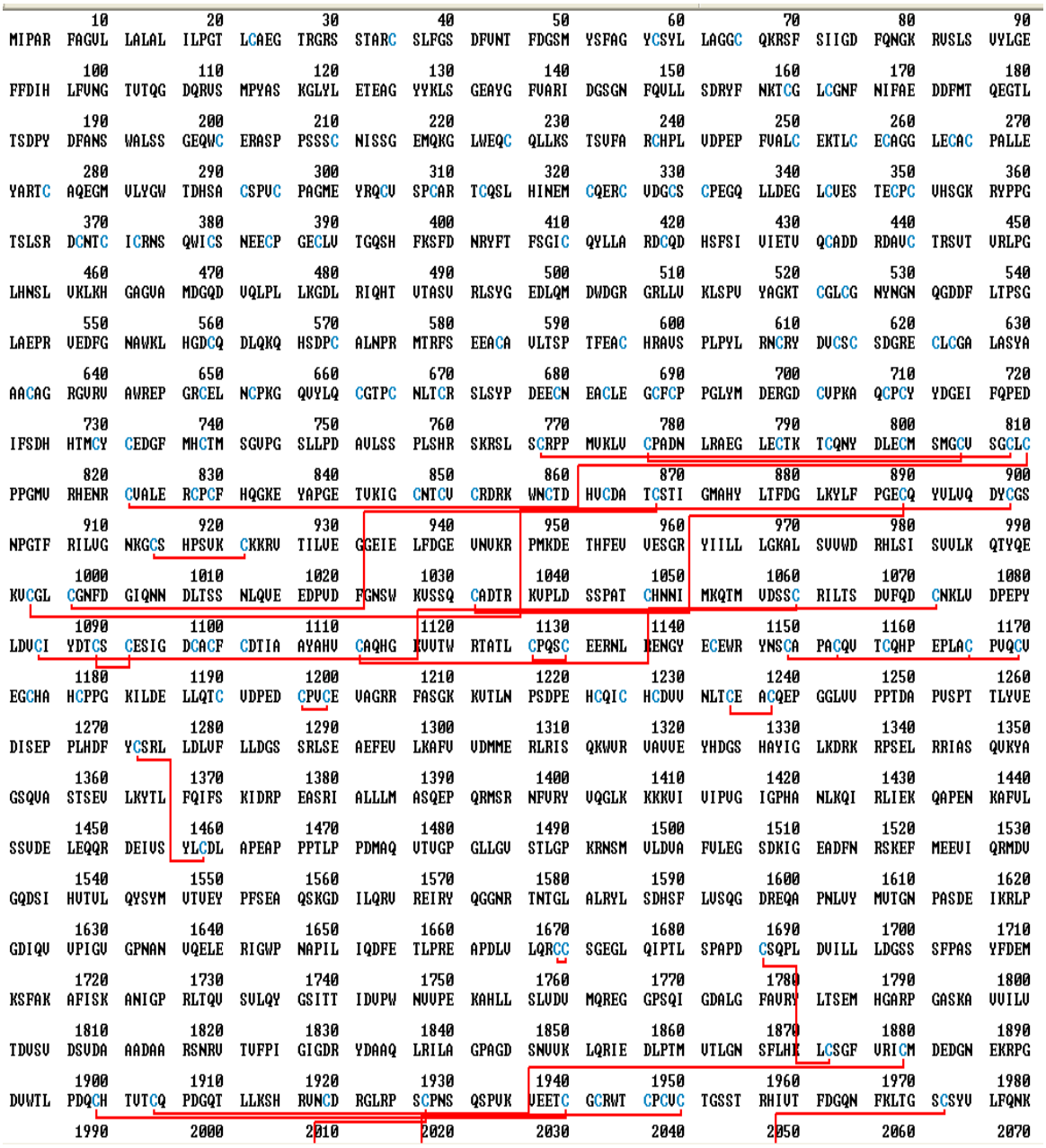
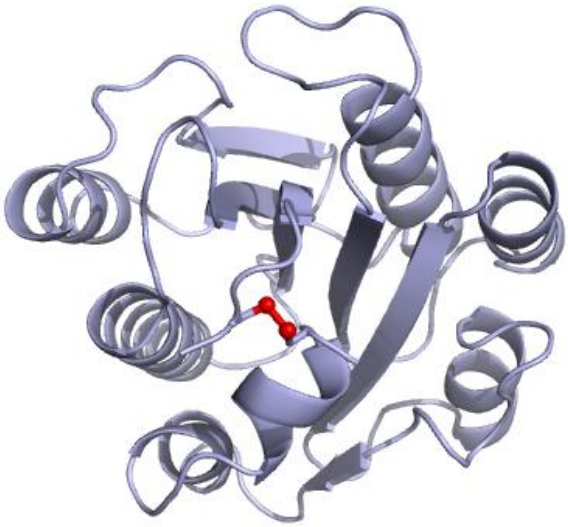


Figure 3.2 VWF A1 domain. From PDB file 1M10 (44).

Figure 3.3 VWF A3 domain. From PDB file 1ATZ (*45*).

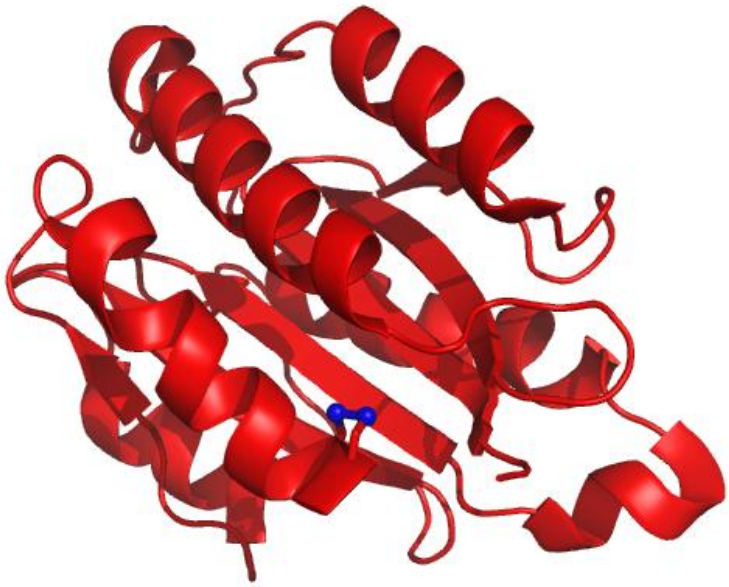


As seen in the crystal structures, each of the VWF A domains has a Rossmann fold with a central β-sheet flanked by α-helices. The A1 domain contains six α-helices; α1, α3, α4, α5, α6, α7 and lacks the α2 helix. Similarly, the A3 domain does not have the α2 helix but has an α8 helix right after the α7 helix. (*45*) The x-ray crystal structure of A1 domain (PDB 1M10) covers 208 residues and defines the sequence from residue 1268 to 1476. The x-ray crystal structure of A2 domain (PDB 3GXB) covers 184 residues and defines the sequence from residue 1495 to 1673. The x-ray crystal structure of A3 domain (PDB 1ATZ) covers 189 residues and defines the sequence from residue 1688 to 1877. Between the A1 and A2 domains 19 residues are not covered in the crystal

structures and between the A2 and A3 domains 15 residues are not covered in the crystal structures.

The A1 and A3 domains contain disulfide bonds linking their N- and C-terminal. The A2 domain contains two vicinal residues (Cys1669-Cys1670), that are buried inside a hydrophobic pocket (*46*). The peptide backbone of Cys1669–Cys1670 and the vicinal disulfide forms an eight-membered ring as shown below in Figure 3.4.

Figure 3.4 VWF A2 domain. From PDB file 3GXB (*15*).



The vital importance of these two cysteines to the function of von Willebrand factor was recently demonstrated. The VWF A2 domain with mutated Cys1669 and Cys1670 and VWF A2 domain with intact vicinal Cys1669Cys1670 were compared in ADAMTS13 proteolysis and circular dichorsim experiments. In this study recombinant A2 domain, covering the residue 1473-1668 fragment without the vicinal disulfides and the residue 1473-1670 fragment with the vicinal disulfides was expressed in

HEK293EBNA cells. According to Luken *et. al* (*15, 46*) the vicinal disulfide bond mainly influences the initial coupling of the α6 helix and primarily regulates the exposure of the ADAMTS-13 binding site. Consequently a second unfolding step is required for the separation of the β5 and α4 loops, since the cleavage site lies in the β4 strand. The results of experiments demonstrated that the A2 domain with native vicinal cysteines was resistant to proteolysis and thermal unfolding. Mutagenesis of Cys1669 and Cys1670 enhanced unfolding and increased susceptibility to ADAMTS13 proteolysis confirming the important role of the vicinal cysteine residues in VWF A2 domain stabilization (*46*).

MATERIALS AND METHODS

Human VWF was purchased from Haematologic Technologies. Glutathione, both reduced (GSH) and oxidized (GSSG), were purchased from Sigma Aldrich and were freshly prepared at a concentration of 10 mM in phosphate buffered saline (PBS; 1.46 mM KH_2PO_4, 9.9 mM Na_2HPO_4, 2.68 mM KCl, 137 mM NaCl, pH 7.4) . The experiments were conducted in the absence of any form of glutathione and in two different reducing environments, a ratio of reduced (GSH) to oxidized glutathione (GSSG) of 1/1 and of 1/1000. For the 1/1 GSH/GSSG ratio, 10 µg deglycosylated protein was incubated with 3.3µl from each of 10mM GSH and GSSG for 2h at 37°C. For the GSH/GSSG 1/1000 ratio, GSH was held constant while GSSH concentration reduced by a factor of a thousand. After incubation, samples were flash frozen using liquid nitrogen and immediately stored at -80°C. In order to minimize the possibility of disulfide scrambling the digestion was conducted under acidic conditions (*47*) in ammonium acetate buffer, pH 3.5. The protein was digested according to manufacturer's instructions (Roche Applied Science, GmbH), using 5 µg sequencing grade endoproteinase GluC in 50mM ammonium acetate buffer at pH 3.5 at room temperature for 24h.

Horse myoglobin protein was purchased from Sigma-Aldrich. Ten micrograms of protein was digested for 24h using 5 µg of sequencing grade endoproteinase GluC in 50mM ammonium acetate buffer (pH 3.5) at room temperature according to manufacturer's instructions.

The GluC digests were then desalted and concentrated using Ziptip C18 pipette tips (Millipore, Bedford, MA) according to the manufacturer's suggested procedure. Two μl of each sample was mixed with an equal volume of 1 M 2,5-dihydroxybenzoic acid (DHB) in 90% methanol containing 0.1% formic acid and spotted onto a Bruker MTP 384 stainless steel MALDI target. The masses of the peptides resulting from GluC digestion were analyzed using MALDI-TOF-MS. Analysis was performed in the mass range of 1-20 kDa using a Bruker Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonik GMBH, Bremen, Germany) in its positive ion mode. MS/MS measurements were performed using its LIFT-TOF/TOF mode.

DATA ANALYSIS

In silico GluC digest fragment ions were generated using Bruker Daltonics sequence editor and exported to the Bruker BioTools 3.1 software to search for the predicted m/z ions in the GluC MALDI-TOF-MS finger print. All the expected ions were searched automatically assuming no more than one Da mass error and no more than two missed cleavages.

RESULTS AND DISSCUSION

In this study, evidence for disulfide rearrangements within the A1A2A3 domains differing from those found in the crystal structures was sought under three different reducing conditions. There are a variety of methods that have been used to determine disulfide bonding patterns in proteins (*48*), however few of them are aimed at discovering and identifying unusual or rare disulfide bonds in a background of more commonly

occurring disulfides. One method based on HPLC separation of these rare conformational isomers has been used, (*49*) but such a technique will not work with the heterogeneous intermolecularly crosslinked von Willebrand factor.

Accordingly, we decided to develop a method wherein von Willebrand factor or any other protein of interest would be cleaved in a specific fashion and the assortment of resulting peptides examined for fragments that correspond in mass to those expected from various peptide disulfide pairs. An important issue is the minimization of possible disulfide scrambling during the cleavage and mass spectrometry measurements. If disulfide scrambling occurs then apparently novel disulfides bonds may very well have no physiological relevance. An interesting solution to this problem was implemented by Moulaei *et al.* (*50*) who used the protease papain under acidic conditions to prevent scrambling when mapping the disulfide topology of scytovirin. Scytovirin is only 95 amino acids long, so the relatively low selectivity of papain was not a major issue. Von Willebrand factor, at over 2000 amino acids long, promised to be far more challenging in finding and correctly identifying peptide pairs made among the 169 cysteines.

We therefore chose to use the endoprotease GluC which, as the name suggests, primarily cleaves on the C-terminal side of glutamate residues. It functions very well at low pH, conditions where disulfide scrambling is minimal. While cleavage at glutamic acid residues is strongly favored by GluC in all buffer conditions, there is the complicating factor that GluC cleaves at aspartic acid residues at a rate 300 times slower than the glutamic acid residues (*51*). The specificity of GluC is also dependent on the structure around the potential cleavage site, for example no cleavage will occur next to a

proline residue. The apparent specificity for glutamate was observed in our experiments in missed cleaved sites at aspartic acid residues. Missed cleavages are the allowed number of targeted amino acids sites per peptide that were not cut, in other words missed cleavage is an error allowance for enzyme inefficiency. A complete digest has zero missed cleavages, means that all sites were cleaved as theoretically expected. In a predicted digest with two missed cleavage all combinations are computed for zero, one, or two uncleaved sites.

Myoglobin was used to establish the efficiency of the GluC digestion at pH ~ 3.5 under identical conditions used for the VWF digestion. This would also enable us to determine how to best treat cleavage at aspartate versus glutamate. Theoretical GluC fragments generated from the myoglobin amino acid sequence were searched for in the MALDI-TOF MS data. MALDI-TOF-MS (Figure 3.5) had about 50% sequence coverage (Fig 3.6). Fragments with two missed cleavages at either an aspartate or a glutamate were observed (Table 3.1). Based on this data we set the maximum missed cleavage threshold as two, and the maximum mass error as 0.4 Da for the data analysis of the VWF experiments.

Figure 3.5: MALDI-TOF-MS finger print obtained from the test GluC digestion of the bovine myoglobin.

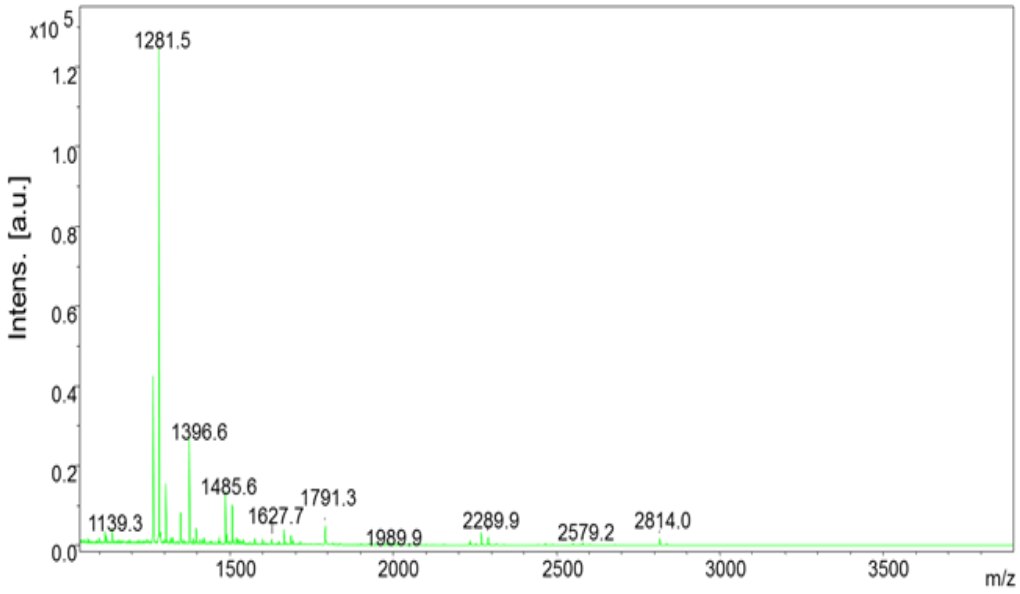


Figure 3.6 : The observed sequence coverage for myoglobin after the GluC digestion at pH 3.5. Peptides identified in the MS data are highlighted in red.

GLSDGEWQLVLNAWGKVEADVAGHGQEVLIRLFTGHPETLEKFDKFKHLKTEAEMKASE
DLKKHGNTVLTALGGILKKKGHHEAEVKHLAESHANKHKIPVKYLEFISDAIIHVLHAKHPSD
FGADAQAAMSKALELFRNDMAAQYKVLGFHG

Table 3.1 : The GluC digest sequences with masses corresponding to the observed fragments in the MALDI-TOF-MS data shown in Fig 3.5.

m/z Theory	m/z Exp.	Error ppm	Missed cleavage	number of D	Sequence
1442.78	1442.65	91.49	0	0	WQLVLNAWGKVE
1420.80	1420.67	85.87	1	1	KFDKFHLKTE
1664.87	1664.84	14.42	1	0	AQAAMSKALELFRND
1281.73	1281.58	121.71	0	0	VLIRLFTGHPE
2290.16	2290.00	72.48	2	2	FISDAIIHVLHAKHPSDFGAD
1409.67	1409.65	18.44	2	0	FGAGAQAAMSKALEAQAAMSKALELFRND

Turning our attention to VWF, based on the expected and observed mass list of the GluC enzymatic fragments listed in Table 3.2, 59% sequence coverage for the entire VWF protein was found when 0.4 Da error and two missed cleavages were considered. The sequence coverage for the A1A2A3 domain was 70%. Sequence coverage of VWF was comparable with myoglobin. Very few peaks were observed in common between the VWF GluC digest fingerprint and GluC blank control, which contained everything except the VWF protein under identical experimental conditions. Fig 3.9 indicates that the majority of the peaks in the VWF GluC fingerprint are due to VWF.

Figure 3.7. MALDI-TOF-MS finger prints of VWF GluC digests under three different reducing conditions.

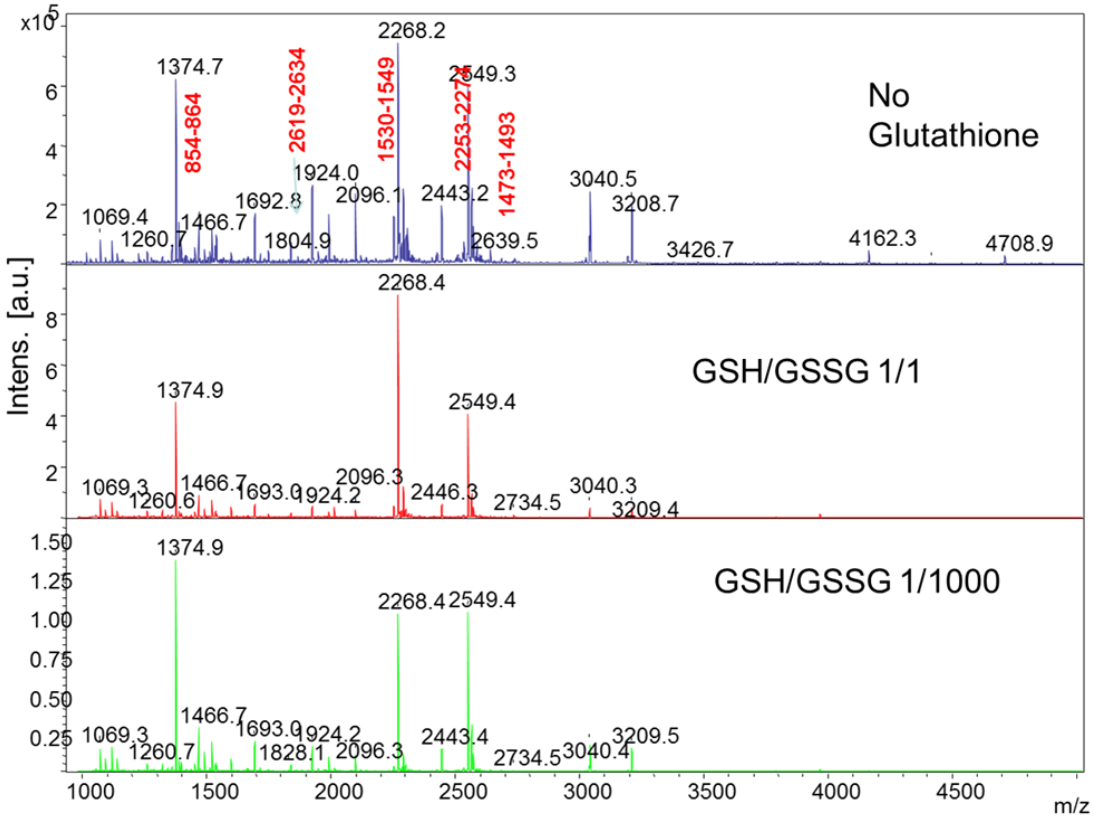


Table 3.2. Representative observed mass list of von Willebrand factor fragments matching predicted fragments within 0.4 Da. This example is the mass list resulting from 1:1 oxidized:reduced glutathione GluC digest. A) Masses corresponding to fully reduced peptides. B) Masses corresponding to two peptides linked by a disulfide.

A)

M/z	Calc. MH+	Error	Dev.(ppm)	Range	Sequence
976.40	976.44	0.04	40.2	2673 - 2680	THFCKVNE
1014.36	1014.32	0.04	37.85	2450 - 2458	VCTCTDMED
1030.35	1030.32	0.03	29.91	1093 - 1102	SIGDCACFCD
1036.57	1037.5	0.09	60.85	677 - 685	EECNEACLE
1073.37	1073.4	0.03	27.06	788 - 796	CTKTCQNYD
1118.60	1118.55	0.05	50.55	2688 - 2697	KRVTGCPPFD
1132.54	1132.55	0.01	8.55	2040 - 2049	VNVYGAIMHE
1248.65	1248.61	0.08	61.99	328 - 339	GCSCPEGQLLDE
1260.67	1260.55	0.12	93.37	1185 - 1195	ELLQTCVDPED
1281.67	1281.38	0.36	247.58	1295 - 1305	VLKAFVVDMME
1282.62	1282.70	0.09	66.02	468 - 479	GQDVQLPLLKGD
1282.62	1282.68	0.06	46.39	2342 - 2353	RGLQPTLTNPGE
1289.66	1289.62	0.04	31.4	1523 - 1533	EVIQRMDVGQD
1319.65	1319.61	0.04	30.38	2241 - 2252	KVMLEGSCVPEE
1374.72	1374.57	0.15	111.53	854 - 864	RKWNCTDHVCD
1386.77	1386.54	0.23	165.81	701 - 712	CVPKAQCPCYYD
1396.74	1396.63	0.11	78.18	713 - 724	GEIFQPEDIFSD
1397.74	1397.64	0.10	96.86	2476 - 2487	SCRSGFTYVLHE
1435.69	1435.78	0.09	65.25	1278 - 1290	LVFLLDGSSRLSE
1450.68	1450.74	0.06	42.68	1435 - 1447	NKAFVLSSVDELE
1510.76	1510.65	0.12	77.55	1132 - 1143	ERNLRENGYECE
1551.74	1551.44	0.29	187.87	1088 - 1102	TCSCESIGDCACFCD
1583.96	1583.70	0.02	154.83	2476 - 2489	SCRSGFTYVLHEGE
1647.85	1647.85	0.00	2.26	936 - 949	LFDGEVNVKRPMKD
1660.73	1660.68	0.05	29.22	1186 - 1200	LLQTCVDPEDCPVCE

M/z	Calc. MH+	Error	Dev.(ppm)	Range	Sequence
976.40	976.44	0.04	40.2	2673 - 2680	THFCKVNE
1714.81	1714.98	0.17	96.31	1360 - 1373	VLKYTLFQIFSKID
1714.81	1714.69	0.12	71.84	698 - 712	RGDCVPKAQCPCYYD
1730.00	1729.83	0.17	85.75	941 - 954	VNVKRPMKDETHFE
1761.93	1761.9	0.03	14.51	1430 - 1445	KQAPENKAFVLSSVDE
1768.90	1768.94	0.04	20.4	1641 - 1655	RIGWPNAPILIQDFE
1804.84	1804.84	0.00	0.11	1599 - 1615	QAPNLVYMVTGNPASDE
1891.21	1890.97	0.23	122.85	2022 - 2039	VTVNGRLVSVPYVGGNME
1906.11	1905.90	0.21	111.05	2150 - 2166	CHKVLAPATFYAICQQD
1908.00	1907.79	0.21	111.03	2234 - 2251	GCFCPPDKVMLEGSCVPE
1961.98	1961.94	0.04	19.8	1597 - 1614	REQAPNLVYMVTGNPASD
1972.91	1972.89	0.00	0.14	107 - 123	QRVSMPYASKGLYLETE
1978.97	1977.99	0.98	494.86	1435 - 1451	NKAFVLSSVDELEQQRD
1979.85	1979.79	0.24	79.91	124 - 141	AGYYKLSGEAYGFVARID
1985.96	1984.19	0.23	114.57	1656-1673	TLPREAPDLVLQRCCSGE
2012.01	2012.07	0.04	18.39	1639 - 1655	LERIGWPNAPILIQDFE
2023.00	2022.83	0.17	83.79	198 - 216	QWCERASPPSSSCNISSGE
2040.51	2040.69	0.82	400.97	1083 - 1096	VCIYDTCSCESIGD
2049.51	2049.17	0.37	279.5	480 - 497	LRIQHTVTASVRLSYGED
2056.98	2057.05	0.07	32.43	941 - 957	VNVKRPMKDETHFEVVE
2081.14	2081.79	0.64	309.5	320 - 338	MCQERCVDGCSCPEGQLLD
2086.32	2085.97	0.35	166.74	2681 - 2697	RGEYFWEKRVTGCPPFD
2118.06	2117.93	0.13	59.56	178 - 197	GTLTSDPYDFANSWALSSGE
2118.23	2117.93	0.30	140.41	178 - 197	GTLTSDPYDFANSWALSSGE
2140.79	2140.07	0.72	336.34	2096 - 2113	FMLRDGTVTTDWKTLVQE
2151.24	2151.06	0.19	86.4	2020 - 2039	MEVTVNGRLVSVPYVGGNME
2169.32	2169.04	0.28	129.05	1986 - 2006	VILHNGACSPGARQGCMKSIE
2179.10	2180.08	0.98	449.45	1250 - 1269	APVSPTTLYVEDISEPPLHD
2268.18	2268.18	0.00	0.55	1568 - 1587	IRYQGGNRTNTGLALRYLSD

M/z	Calc. MH+	Error	Dev.(ppm)	Range	Sequence
976.40	976.44	0.04	40.2	2673 - 2680	THFCKVNE
2268.18	2268.12	0.06	27.1	1530 - 1549	VGQDSIHVTVLQYSYMVTVE
2268.18	2267.96	0.23	100.53	257 - 278	CAGGLECACPALLEYARTCAQE
2268.36	2267.96	0.31	248.65	257 - 278	CAGGLECACPALLEYARTCAQE
2295.12	2295.15	0.03	12.44	1664 - 1685	LVLQRCCSGEGLQIPTLSPAPD
2332.25	2333.02	0.78	332.88	2699 - 2720	HKCLAEGGKIMKIPGTCCDTCE
2354.13	2354.01	0.12	49.48	2619 - 2639	CRKTTCNPCPLGYKEENNTGE
2413.49	2413.36	0.13	54.87	1616 - 1638	IKRLPGDIQVVPIGVGPNANVQE
2479.21	2479.15	0.05	11.44	25 - 47	GTRGRSSTARCSLFGSDFVNTFD
2505.27	2505.01	0.27	105.68	2222 - 2245	GNVSSCGDHPSEGCFCPPDKVMLE
2534.25	2534.41	0.16	61.77	1833 - 1856	AAQLRILAGPAGDSNVVKLQRIED
2587.20	2587.15	0.05	20.17	1172 - 1195	GCHAHCPPGKILDELLQTCVDPED
2640.49	2640.42	0.07	23.56	1473 - 1498	MAQVTVGPGLLGVSTLGPKRNSMVLD
2675.37	2675.48	0.12	43.79	1616 - 1640	IKRLPGDIQVVPIGVGPNANVQELE
2685.29	2685.30	0.02	6.65	1034 - 1057	TRKVPLDSSPATCHNNIMKQTMVD
2753.32	2754.19	0.87	315.07	173 - 197	FMTQEGTLTSDPYDFANSWALSSGE
2807.38	2808.00	0.62	220.61	2214 - 2240	HGCPRHCDGNVSSCGDHPSEGCFCPPD
3008.53	3008.65	0.12	40.74	471 - 497	VQLPLLKGDLRIQHTVTASVRLSYGED
3061.39	3061.43	0.10	312.3	1588 - 1615	HSFLVSQGDREQAPNLVYMVTGNPASDE
3078.40	3078.48	0.07	23.87	1534 - 1560	SIHVTVLQYSYMVTVEYPFSEAQSKGD
3194.71	3193.73	0.97	305.1	468 - 496	GQDVQLPLLKGDLRIQHTVTASVRLSYGE
3212.58	3212.49	0.08	25.33	142 - 170	GSGNFQVLLSDRYFNKTCGLCGNFNIFAE
3377.36	3377.56	0.21	60.63	2635-2665	NNTGECCGRCLPTACTIQLRGGQIMTLKRDE
3496.86	3495.30	0.05	41	471 - 501	VQLPLLKGDLRIQHTVTASVRLSYGEDLQMD
3522.78	3523.78	0.20	283.69	1568 - 1598	IRYQGGNRTNTGLALRYLSDHSFLVSQGDRE
3524.18	3523.78	0.40	112.48	1568 - 1598	IRYQGGNRTNTGLALRYLSDHSFLVSQGDRE
3572.72	3572.51	0.20	56.25	650 - 681	LNCPKGQVYLQCGTPCNLTCRSLSYPDEECNE
3860.22	3860.79	0.57	146.83	2599 - 2633	VCTTCRCMVQVGVISGFKLECRKTTCNPCPLGYKE
3860.22	3859.75	0.47	121.75	388 - 420	CLVTGQSHFKSFDNRYFTFSGICQYLLARDCQD

M/z	Calc. MH+	Error	Dev.(ppm)	Range	Sequence
976.40	976.44	0.04	40.2	2673 - 2680	THFCKVNE
3861.05	3860.79	0.26	67.11	2599 - 2633	VCTTCRCMVQVGVISGFKLECRKTTCNPCPLGYKE
3923.19	3923.68	0.39	125.39	2488 - 2525	GECCGRCLPSACEVVTGSPRGDSQSSWKSVGSQ WASPE
3943.28	3942.79	0.40	125.5	548 - 581	DFGNAWKLHGDCQDLQKQHSDPCALNPRMTRFSE
4184.21	4184.57	0.36	158.69	583 - 620	ACAVLTSPTFEACHRAVSPLPYLRNCRYDVCSCSDGRE
4184.25	4184.21	0.04	9.45	1464 - 1504	APPPTLPPDMAQVTVGPGLLGVSTLGPKRNSMVLDV AFVLE
4424.58	4424.84	0.26	58.47	2767 - 2803	VQDQCSCCSPTRTEPMQVALHCTNGSVVYHEVLNAME
4595.57	4595.91	0.34	73.05	819 - 840	CMSMGCVSGCLCPPGMVRHENRCVALERCPCFHQGKE
5176.22	5176.38	0.16	30.54	2270 - 2317	AWVPDHQPCQICTCLSGRKVNCTTQPCPTAKAPTCGLCE VARLRQNAD
6008.25	6007.21	0.02	8.99	1377 - 1429	ASRIALLLMASQEPQRMSRNFVRYVQGLKKKKVIVIPVGIGP HANLKQIRLIE

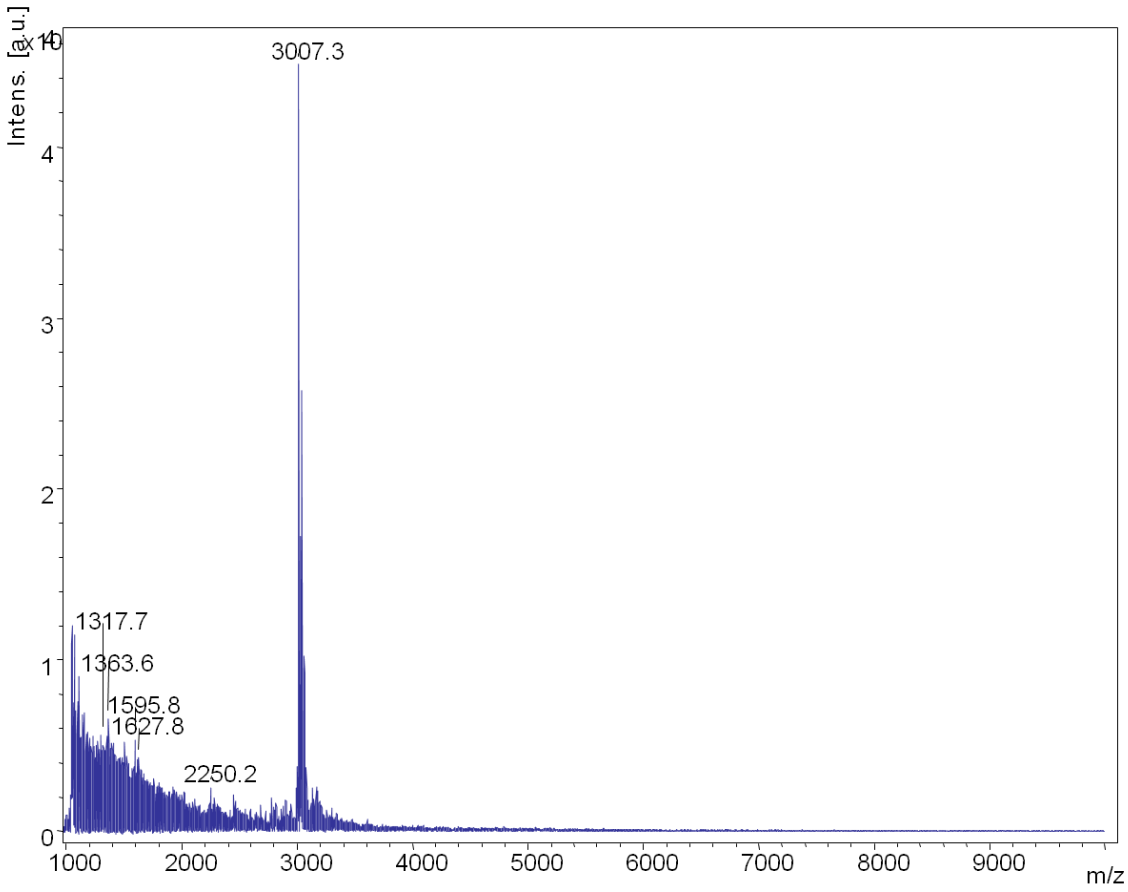
B)

M/z	Calc. MH+	Error	Dev.(ppm)	1st Range	Sequence
3007.44	3007.41	0.02	8.97	1857-1881	LPTMVTLGNSFLHKLCSGFVRICMDE / CSQPLD
3202.62	3202.53	0.09	27.19	1021 - 1040	FGNSWKVSSQCADTRKVPLD / CQYVLVQD
3242.62	3242.61	0.01	3.57	1270 - 1290	FYCSRLLDLVFLLDGSSRLSE / IVSYLCD
3250.54	3250.57	0.03	9.41	1452 - 1472	EIVSYLCDLAPEAPPPTLPPD / FYCSRLLD
3250.54	3250.55	0.01	2.7	1270-1277	EIVSYLCDLAPEAPPPTLPPD / FYCSRLLD
3265.57	3265.46	0.10	31.02	2745 - 2763	VDIHYCQGKCASKAMYSID / CKCSPRKCSK
3265.57	3265.46	0.10	31.02	2745 - 2763	VDIHYCQGKCASKAMYSID / CKCSPRKCSK
3496.71	3499.59	0.28	21.62	1071 - 1092	CNKLVDPEPYLDVCIYDTCSCE / SSCRILTSD
3607.94	3607.62	0.33	90.42	2745 - 2766	VDIHYCQGKCASKAMYSIDIND / CKCSPRKCSK
3682.96	3682.77	0.19	50.27	2721 - 2726	EPECND/ QCSCCSPTRTEPMQVALHCTNGSVVYHE
3735.91	3735.68	0.24	63.66	2747 - 2769	IHYCQGKCASKAMYSIDINDVQD / CKCSPRKCSK
3735.91	3735.68	0.24	63.66	2747 - 2769	IHYCQGKCASKAMYSIDINDVQD / CKCSPRKCSK
4856.38	4856.17	0.04	23.85	2727 - 2742	ITARLQYVKVGSCKSE / QCSCCSPTRTEPMQVALHCTNGSVVYHE
5137.23	5137.39	0.16	30.71	1071 - 1087	CNKLVDPEPYLDVCIYD / TIAAYAHVCAQHGKVVTWRTATLCPQSCE
6008.92	6008.83	0.12	39.8	799 - 825	CMSMGCVSGCLCPPGMVRHENRCVALE / AVLSSPLSHRSKRSLSCRPPMVKLVCPAD
6775.68	6775.47	0.29	95.88	955 - 1000	VVESGRYIILLLGKALSVVWDRHLSISVVLKQTYQE KVCGLCGNFD / ATCSTIGMAHYLTFD
7089.48	7089.57	0.08	11.69	958 - 1006	SGRYIILLLGKALSVVWDRHLSISVVLKQTYQEKV CGLCGNFDGIQNND / ATCSTIGMAHYLTFD
2714.39	2714.28	0.11	39.58	1262-1277	ISEPPLHDFYCSRLLD / IVSYLCD

Figure 3.8. The sequence coverage of the enzymatic peptide mass fingerprint of VWF. The sequence coverage in the A1A2A3 domain is italicized and colored in blue. The sequence coverage falling outside of the A1A2A3 domain colored in red.

```
   1 miparfagvl lalalilpgt lcaegtrgrs starcslfgs dfvntfdgsm ysfagycsyl
  61 laggcqkrsf siigdfqngk rvslsvylge ffdihlfvng tvtqgdqrvs mpyaskglyl
 121 eteagyykls geaygfvari dgsgnfqvll sdryfnktcg lcgnfnifae ddfmtqegtl
 181 tsdpydfans walssgeqwc erasppsssc nissgemqkg lweqcqllks tsvfarchpl
 241 vdpepfvalc ektlcecagg lecacpalle yartcaqegm vlygwtdhsa cspvcpagme
 301 yrqcvspcar tcqslhinem cqercvdgcs cpegqlldeg lcvestecpc vhsgkryppg
 361 tslsrdcntc icrnsqwics neecpgeclv tgqshfksfd nryftfsgic qyllardcqd
 421 hsfsivietv qcaddrdavc trsvtvrlpg lhnslvklkh gagvamdgqd vqlpllkgdl
 481 riqhtvtasv rlsygedlqm dwdgrgrllv klspvyagkt cglcgnyngn qgddfltpsg
 541 laeprvedfg nawklhgdcq dlqkqhsdpc alnprmtrfs eeacavltsp tfeachravs
 601 plpylrncry dvcscsdgre clcgalasya aacagrgvrv awrepgrcel ncpkgqvylq
 661 cgtpcnltcr slsypdeecn eaclegcfcp pglymdergd cvpkaqcpcy ydgeifqped
 721 ifsdhhtmcy cedgfmhctm sgvpgsllpd avlssplshr skrslscrpp mvklvcpadn
 781 lraeglectk tcqnydlecm smgcvsgclc ppgmvrhenr cvalercpcf hqgkeyapge
 841 tvkigcntcv crdrkwnctd hvcdatcsti gmahyltfdg lkylfpgecq yvlvqdycgs
 901 npgtfrilvg nkgcshpsvk ckkrvtilve ggeielfdge vnvkrpmkde thfevvesgr
 961 yiilllgkal svvwdrhlsi svvlkqtyqe kvcglcgnfd giqnndltss nlqveedpvd
1021 fgnswkvssq cadtrkvpld sspatchnni mkqtmvdssc riltsdvfqd cnklvdpepy
1081 ldvciydtcs cesigdcacf cdtiaayahv caqhgkvvtw rtatlcpqsc eernlrengy
1141 ecewrynsca pacqvtcqhp eplacpvqcv egchahcppg kildellqtc vdpedcpvce
1201 vagrrfasgk kvtlnpsdpe hcqichcdvv nltceacqep gglvvpptda pvspttlyve
1261 diseppthdf ycsrlldlvf lldgssrlse aefevlkafv vdmmerlris qkwvrvavve
1321 yhdgshayig lkdrkrpsel rriasqvkya gsqvastsev lkytlfqifs kidrpeasri
1381 alllmasqep qrmsrnfvry vqglkkkkvi vipvgigpha nlkqirliek qapenkafvl
1441 ssvdeleqqr deivsylcdl apeappptlp pdmaqvtvgp gllgvstlgp krnsmvldva
1501 fvlegsdkig eadfnrskef meeviqrmdv gqdsihvtvl qysymvtvey pfseaqskgd
1561 ilqrvreiry qggnrtntgl alrylsdhsf lvsqgdreqa pnlvymvtgn pasdeikrlp
1621 gdiqvvpigv gpnanvqele rigwpnapil iqdfetlpre apdlvlqrcc sgeglqiptl
1681 spapdcsqpl dvillldgss sfpasyfdem ksfakafisk anigprltqv svlqygsitt
1741 idvpwnvvpe kahllslvdv mqreggpsqi gdalgfavry ltsemhgarp gaskavvilv
1801 tdvsvdsvda aadaarsnrv tvfpigigdr ydaaqlrila gpagdsnvvk lqriedlptm
1861 vtlgnsflhk lcsgfvricm dedgnekrpg dvwtlpdqch tvtcqpdgqt llkshrvncd
1921 rglrpscpns qspvkveetc gcrwtcpcvc tgsstrhivt fdgqnfkltg scsyvlfqnk
1981 eqdlevilhn gacspgarqg cmksievkhs alsvelhsdm evtvngrlvs vpyvggnmev
2041 nvygaimhev rfnhlghift ftpqnnefql qlspktfask tyglcgicde ngandfmlrd
2101 gtvttdwktl vqewtvqrpg qtcqpileeq clvpdsshcq vlllplfaec hkvlapatfy
2161 aicqqdschq eqvceviasy ahlcrtngvc vdwrtpdfca mscppslvyn hcehgcprhc
2221 dgnvsscgdh psegcfcppd kvmlegscvp eeactqcige dgvqhqflea wvpdhqpcqi
2281 ctclsgrkvn cttqpcptak aptcglceva rlrqnadqcc peyecvcdpv scdlppvphc
2341 erglqptltn pgecrpnftc acrkeeckrv sppscpphrl ptlrktqccd eyecacncvn
2401 stvscplgyl astatndcgc ttttclpdkv cvhrstiypv gqfweegcdv ctctdmedav
2461 mglrvaqcsq kpcedscrsg ftyvlhegec cgrclpsace vvtgsprgds qsswksvgsq
2521 waspenpcli necvrvkeev fiqqrnvscp qlevpvcpsg fqlscktsac cpscrcerme
2581 acmlngtvig pgktvmidvc ttcrcmvqvg visgfklecr kttcnpcplg ykeennтgec
2641 cgrclptact iqlrggqimt lkrdetlqdg cdthfckvne rgeyfwekrv tgcppfdehk
2701 claeggkimk ipgtccdtce epecnditar lqyvkvgsck sevevdihyc qgkcaskamy
2761 sidindvqdq csccsptrte pmqvalhctn gsvvyhevln ameckcsprk csk
```

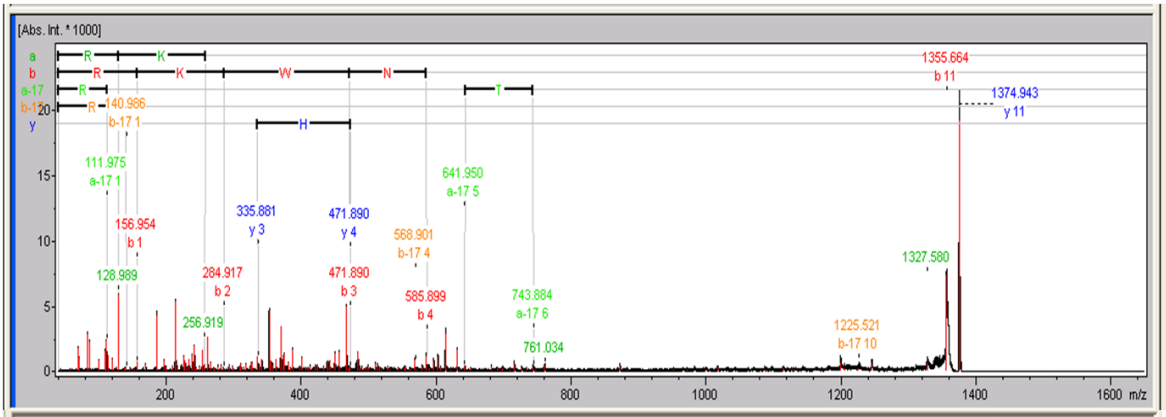
Figure 3.9: MALDI-TOF-MS for control “blank” which contains everything but VWF under identical conditions.



In order to further confirm and verify their identity to the VWF sequence, two high intensity peaks at m/z 1374 and 1117 in the GluC peptide mass fingerprint were analyzed further by using MALDI LIFT-TOF/TOF (MS/MS). Typical fragment ions expected in LIFT-TOF/TOF measurements are “b” and “a” series ions (N-terminus fragments) and “y” series ions (C-terminus fragments). According to the VWF sequence, fragment peaks at 1374 and 1117 in the MALDI-TOF-MS finger print in Figure 3.10 and Figure 3.11 correspond to the predicted peptides (Table 3.2) RKWNCTDHVCD and

KRVTGCPPFD. Figure 3.10 shows how well the observed fragment ions match with the expected “a”, “b” and “y” ions from the assigned sequences.

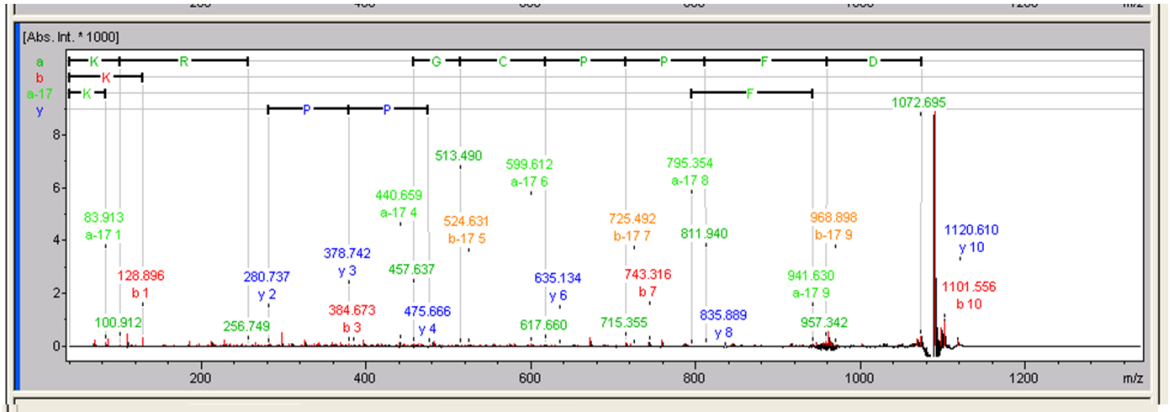
Figure 3.10: MALDI LIFT-TOF/TOF spectra obtained for the peak m/z 1374 in the VWF GluC finger print in figure x. Theoretical and observed masses of “a”, “b” and “y” ions with respect to assigned sequence RKWNCTDHVCD in Table 3.2 are shown below.



	R	K	W	N	C	T	D	H	V	C	D	Arg	Lys	Trp	Asn	Cys	Thr	Asp	His	Val	Cys	Asp
Ion	1	2	3	4	5	6	7	8	9	10	11	1		2		3		4		5		6		7		8		9		10		11	
a	R	K	W	N	C	T	D	H	V	C	D	129.113	257.208	443.288	557.331	659.332	760.380	875.407	1012.466	1111.534	1213.535	1328.562
b	R	K	W	N	C	T	D	H	V	C	D	157.108	285.203	471.283	585.326	687.327	788.375	903.402	1040.460	1139.529	1241.530	1356.557
a-17	R	K	W	N	C	T	D	H	V	C	D	112.087	240.182	426.261	540.304	642.305	743.353	858.380	995.439	1094.507	1196.509	1311.536
b-17	R	K	W	N	C	T	D	H	V	C	D	140.082	268.177	454.256	568.299	670.300	771.348	886.375	1023.434	1122.502	1224.504	1339.531
y	R	K	W	N	C	T	D	H	V	C	D	134.045	236.046	335.115	472.173	587.200	688.248	790.249	904.292	1090.372	1218.467	1374.568
i	R	K	W	N	C	T	D	H	V	C	D	129.113	101.107	159.091	87.055	75.013	74.059	88.039	110.071	72.080	75.013	88.039
	11	10	9	8	7	6	5	4	3	2	1	Asp	Cys	Val	His	Asp	Thr	Cys	Asn	Trp	Lys	Arg

The data confirms that the identity of VWF. The correspondence between the expected and observed data also suggests that the disulfide rearrangements of the A1A2A3 domain could likewise be probed based on the same sort of experimental data. (Dr. Jack Lay personal communication).

Figure 3.11: MALDI LIFT-TOF/TOF spectra obtained for the peak m/z 1118 in the VWF GluC finger print in Figure 3.7 Theoretical and observed masses of "a", "b" and "y" ions with respect to assigned sequence KRVTGCPPFD in Table 3.2 are shown below.



	K	R	V	T	G	C	P	P	F	D	Lys	Arg	Val	Thr	Gly	Cys	Pro	Pro	Phe	Asp
Ion	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
a	K	R	V	T	G	C	P	P	F	D	101.107	257.208	356.277	457.325	514.346	617.355	714.408	811.461	958.529	1073.556
b	K	R	V	T	G	C	P	P	F	D	129.102	285.203	384.272	485.319	542.341	645.350	742.403	839.456	986.524	1101.551
a-17	K	R	V	T	G	C	P	P	F	D	84.081	240.182	339.250	440.298	497.319	600.329	697.381	794.434	941.503	1056.530
b-17	K	R	V	T	G	C	P	P	F	D	112.076	268.177	367.245	468.293	525.314	628.324	725.376	822.429	969.497	1084.524
y	K	R	V	T	G	C	P	P	F	D	134.045	281.113	378.166	475.219	578.228	635.249	736.297	835.365	991.467	1119.562
i	K	R	V	T	G	C	P	P	F	D	101.107	129.113	72.080	74.059	30.033	76.021	70.065	70.065	120.080	88.039
	10	9	8	7	6	5	4	3	2	1	Asp	Phe	Pro	Pro	Cys	Gly	Thr	Val	Arg	Lys

One hundred and two possible GluC digest fragments (and their expected m/z values) corresponding to possible disulfide arrangements entirely within the A1A2A3 domain were generated with Bruker Daltonics Sequence Editor 3.1 software and are shown in Table 3.3. The MALDI mass spectral data was then searched for the 102 possible ions using the Bruker Daltonics software with the minimum mass error it allows, one Dalton, and assuming no more than two missed cleavages. Some of these are less likely than others as they assume missed cleavages at glutamate.

Table 3.3. Predicted GluC enzymatic fragments constructed from possible disulfide arrangements in A1A2A3 domain. Entries that are italicized contain one or more missed cleavages at a glutamate. Crosslinked peptides 1-27 are those that are predicted on the basis of previously assigned disulfide bonding partners. Crosslinked peptide 28-102 are those that would not be expected if only the literature arrangement were present.

	Disulfide Bond	**Mass**	**Sequence**
1	Cys1272-Cys1458	1825.9	FYCSRLLD/IVSYLCD
2	*Cys1272-Cys1458*	*1954.9*	*FYCSRLLD/EIVSYLCD*
3	Cys1272-Cys1458	2236.1	FYCSRLLD/IVSYLCDLAPE
4	*Cys1272-Cys1458*	*2365.1*	*FYCSRLLD/EIVSYLCDLAPE*
5	*Cys1272-Cys1458*	*2482.1*	*FYCSRLLD/QQRDEIVSYLCD*
6	*Cys1272-Cys1458*	*2655.3*	*FYCSRLLDLVFLLD/EIVSYLCD*
7	*Cys1272-Cys1458*	*2714.3*	*ISEPPLHDFYCSRLLD/IVSYLCD*
8	*Cys1272-Cys1458*	*2829.3*	*DISEPPLHD/FYCSRLLDEIVSYLCD*
9	*Cys1272-Cys1458*	*2843.3*	*ISEPPLHDFYCSRLLD/EIVSYLCD*
10	Cys1272-Cys1458	2936.5	FYCSRLLDLVFLLD/IVSYLCDLAPE
11	Cys1272-Cys1458	3124.5	ISEPPLHDFYCSRLLD/IVSYLCDLAPE
12	Cys1272-Cys1458	3242.6	FYCSRLLDLVFLLDGSSRLSE/IVSYLCD
13	*Cys1272-Cys1458*	*3250.6*	*EIVSYLCDLAPEAPPPTLPPD/FYCSRLLD*
14	Cys1686-Cys1872	3007.4	CSQPLD/LPTMVTLGNSFLHKLCSFGVRICMD[§]
15	Cys1686-Cys1872	4105.1	GLQIPTLSPAPDCSQPLD/LPTMVTLGNSFLHKLCSFGVRICMD
16	Cys1686-Cys1872	3489.8	CSQPLDVILLLD/DLPTMVTLGNSFLHKLCSFGVRICMD
17	Cys1686-Cys1872	3604.9	CSQPLDVILLLD/LPTMVTLGNSFLHKLCSFGVRICMD
18	Cys1686-Cys1872	4013.0	GLQIPTLSPAPDCSQPLD/DLPTMVTLGNSFLHKLCSFGVRICMD
19	Cys1686-Cys1872	4128.1	GLQIPTLSPAPDCSQPLD/LPTMVTLGNSFLHKLCSFGVRICMD
20	Cys1686-Cys1872	4635.3	CSQPLDVILLLDGSSSFPASYFD/DLPTMVTLGNSFLHKLCSFGVRICMD
21	Cys1686-Cys1872	4679.5	GLQIPTLSPAPDCSQPLDVILLLD/DLPTMVTLGNSFLHKLCSFGVRICMD
22	*Cys1669-Cys1670*[‡]	*5100.5*	*LVLQRCCSGEGLQIPTLSPAPDCSQPLD/LPTMVTLGNSFLHKLCSGFV*
23	Cys1669-Cys1670[†]	1388.6	APDLVLQRCCSGE
24	*Cys1669-Cys1670*[†]	*2578.2*	*APDLVLQRCCSGEGLQIPTLSPAPD*
25	Cys1669-Cys1670[†]	1105.5	LVLQRCCSGE
26	*Cys1669-Cys1670*[†]	*2295.1*	*LVLQRCCSGEGLQIPTLSPAPD*
27	*Cys1669-Cys1670*[†]	*1985.9*	*TLPREAPDLVLQRCCSGE*
28	*Cys1458-VC**	*1915.9*	*QQRDEIVSYLCD/LVLQRCCSGEGLQIPTLSPAPD*
29	*Cys1458-VC**	*2044.9*	*EIVSYLCD/LVLQRCCSGE*
30	*Cys1458-VC**	*2199.0*	*QQRDEIVSYLCD/APDLVLQRCCSGE*
31	Cys1458-VC*	2326.1	IVSYLCDLAPE/LVLQRCCSGE
32	*Cys1458-VC**	*2328.1*	*EIVSYLCD/APDLVLQRCCSGE*

33	*Cys1458-VC**	*2455.2*	*EIVSYLCDLAPE/LVLQRCCSGE*
34	*Cys1458-VC**	*2571.2*	*QQRDEIVSYLCD/LVLQRCCSGE*
35	Cys1458-VC*	2609.2	IVSYLCDLAPE/APDLVLQRCCSGE
36	*Cys1458-VC**	*2795.3*	*QQRDEIVSYLCD/TLPREAPDLVLQRCCSGE*
37	*Cys1458-VC**	*3105.5*	*QQRDEIVSYLCD/LVLQRCCSGEGLQIPTLSPAPD*
38	*Cys1458-VC**	*3234.6*	*EIVSYLCD/LVLQRCCSGEGLQIPTLSPAPD*
39	*Cys1458-VC**	*3387.6*	*IVSYLCD/APDLVLQRCCSGEGLQIPTLSPAPD*
40	*Cys1458-VC**	*3515.7*	*IVSYLCDLAPE/APDLVLQRCCSGEGLQIPTLSPAPD*
41	*Cys1458-VC**	*3746.8*	*QQRDEIVSYLCD/LVLQRCCSGEGLQIPTLSPAPDCSQPLD*
42	Cys1272-VC*	2120.0	FYCSRLLD/LVLQRCCSGE
43	Cys1272-VC*	2403.1	FYCSRLLD/APDLVLQRCCSGE
44	Cys1272-VC*	2820.4	FYCSRLLDLVFLLD/LVLQRCCSGE
45	*Cys1272-VC**	*2999.4*	*FYCSRLLD/TLPREAPDLVLQRCCSGE*
46	*Cys1272-VC**	*3008.4*	*ISEPPLHDFYCSRLLD/LVLQRCCSGE*
47	Cys1272 -VC*	3103.5	FYCSRLLDLVFLLD/APDLVLQRCCSGE
48	*Cys1272-VC**	*3123.5*	*DISEPPLHDFYCSRLLD/LVLQRCCSGE*
49	*Cys1272-VC**	*3291.5*	*ISEPPLHDFYCSRLLD/APDLVLQRCCSGE*
50	Cys1272-VC*	3536.8	FYCSRLLDLVFLLDGSSRLSE/LVLQRCCSGE
51	*Cys1272-VC**	*3592.7*	*FYCSRLLD/APDLVLQRCCSGEGLQIPTLSPAPD*
52	*Cys1272-VC**	*3708.8*	*ISEPPLHDFYCSRLLDLVFLLD/LVLQRCCSGE*
53	*Cys1272-VC**	*3951.9*	*FYCSRLLD/LVLQRCCSGEGLQIPTLSPAPDCSQPLD*
54	*Cys1272-VC**	*4198.1*	*ISEPPLHDFYCSRLLD/LVLQRCCSGEGLQIPTLSPAPD*
55	Cys1686-VC*	2048.9	CSQPLD/APDLVLQRCCSGE
56	Cys1686-VC*	2432.2	CSQPLDVILLLD/LVLQRCCSGE
57	*Cys1686-VC**	*2645.2*	*CSQPLD/TLPREAPDLVLQRCCSGE*
58	Cys1686-VC*	2715.3	CSQPLDVILLLD/APDLVLQRCCSGE
59	Cys1686-VC*	2955.4	GLQIPTLSPAPDCSQPLD/LVLQRCCSGE
60	Cys1686-VC*	3238.5	GLQIPTLSPAPDCSQPLD/APDLVLQRCCSGE
61	Cys1686-VC*	3577.7	CSQPLDVILLLDGSSSFPASYFD/LVLQRCCSGE
62	Cys1686-VC*	3621.9	GLQIPTLSPAPDCSQPLDVILLLD/LVLQRCCSGE
63	VC-Cys1872*	3267.6	LVLQRCCSGE/DLPTMVTLGNSFLHKLCSFGVRICMD
64	VC-Cys1872*	3382.7	LVLQRCCSGE/LPTMVTLGNSFLHKLCSFGVRICMD
65	VC-Cys1872*	3550.8	APDLVLQRCCSGE/DLPTMVTLGNSFLHKLCSFGVRICMD
66	VC-Cys1872*	3665.8	APDLVLQRCCSGE/LPTMVTLGNSFLHKLCSFGVRICMD
67	*VC-Cys1872**	*4147.1*	*TLPREAPDLVLQRCCSGE/DLPTMVTLGNSFLHKLCSFGVRICMD*
68	*VC-Cys1872**	*4457.3*	*DLPTMVTLGNSFLHKLCSFGVRICMD/ LVLQRCCSGEGLQIPTLSPAPD*
69	*VC-Cys1872**	*4572.3*	*LVLQRCCSGEGLQIPTLSPAPD/LPTMVTLGNSFLHKLCSFGVRICMD*
70	*VC-Cys1872**	*4740.4*	*APDLVLQRCCSGEGLQIPTLSPAPD/DLPTMVTLGNSFLHKLCSFGVRICMD*

71	*VC-Cys1872**	*5099.5*	*LVLQRCCSGEGLQIPTLSPAPDCSQPLD/DLPTMVTLGNSFLHKLCSFGVRICMD*
72	Cys1458-Cys1686	1471.7	IVSYLCD/CSQPLD
73	*Cys1458-Cys1686*	*1600.7*	*EIVSYLCD/CSQPLD*
74	*Cys1458-Cys1686*	*2010.9*	*EIVSYLCDLAPE/CSQPLD*
75	*Cys1458-Cys1686*	*2127.9*	*QQRDEIVSYLCD/CSQPLD*
76	Cys1458-Cys1686	2138.1	IVSYLCD/CSQPLDVILLLD
77	*Cys1458-Cys1686*	*2267.1*	*EIVSYLCD/CSQPLDVILLLD*
78	Cys1458-Cys1686	2548.3	IVSYLCDLAPE/CSQPLDVILLLD
79	Cys1458-Cys1686	2661.3	IVSYLCD/GLQIPTLSPAPDCSQPLD
80	*Cys1458-Cys1686*	*2790.3*	*EIVSYLCD/CSQPLD*
81	Cys1458-Cys1686	3071.5	IVSYLCDLAPE/GLQIPTLSPAPDCSQPLD
82	Cys1458-Cys1686	3327.7	IVSYLCD/GLQIPTLSPAPDCSQPLDVILLLD
83	*Cys1458-Cys1686*	*3747.8*	*IVSYLCD/LVLQRCCSGEGLQIPTLSPAPDCSQPLD*
84	Cys1458-Cys1686	1881.8	IVSYLCDLAPE/CSQPLD
85	Cys1272-Cys1872	3177.6	FYCSRLLD/LPTMVTLGNSFLHKLCSFGVRICMD
86	Cys1272-Cys1872	3292.6	FYCSRLLD/LPTMVTLGNSFLHKLCSFGVRICMD
87	Cys1272-Cys1872	3878.0	FYCSRLLDLVFLLD/DLPTMVTLGNSFLHKLCSFGVRICMD
88	Cys1272-Cys1872	3993.0	FYCSRLLDLVFLLD/LPTMVTLGNSFLHKLCSFGVRICMD
89	*Cys1272-Cys1872*	*4066.0*	*ISEPPLHDFYCSRLLD/DLPTMVTLGNSFLHKLCSFGVRICMD*
90	*Cys1272-Cys1872*	*4181.1*	*ISEPPLHDFYCSRLLD/LPTMVTLGNSFLHKLCSFGVRICMD*
91	Cys1272-Cys1872	4459.3	FYCSRLLD/LPTMVTLGNSFLHKLCSFGVRICMD
92	Cys1272-Cys1872	4594.4	FYCSRLLDLVFLLDGSSRLSE/DLPTMVTLGNSFLHKLCSFGVRICMD
93	*Cys1272-Cys1872*	*4766.5*	*ISEPPLHDFYCSRLLDLVFLLD/DLPTMVTLGNSFLHKLCSFGVRICMD*
94	Cys1458-Cys1872	2973.5	IVSYLCD/DLPTMVTLGNSFLHKLCSFGVRICMD
95	Cys1458-Cys1872	3088.5	IVSYLCD/DLPTMVTLGNSFLHKLCSFGVRICMD
96	*Cys1458-Cys1872*	*3102.5*	*EIVSYLCD/DLPTMVTLGNSFLHKLCSFGVRICMD*
97	*Cys1458-Cys1872*	*3217.6*	*EIVSYLCD/LPTMVTLGNSFLHKLCSFGVRICMD*
98	Cys1458-Cys1872	3383.7	IVSYLCDLAPE/DLPTMVTLGNSFLHKLCSFGVRICMD
99	Cys1458-Cys1872	3498.7	IVSYLCDLAPE/LPTMVTLGNSFLHKLCSFGVRICMD
100	*Cys1458-Cys1872*	*3512.8*	*EIVSYLCDLAPE/DLPTMVTLGNSFLHKLCSFGVRICMD*
101	*Cys1458-Cys1872*	*3629.8*	*QQRDEIVSYLCD/DLPTMVTLGNSFLHKLCSFGVRICMD*
102	Cys1458-Cys1872	4255.2	IVSYLCD/LPTMVTLGNSFLHKLCSFGVRICMD

[§]A peak of mass 3007.3 is also observed in the blank mass spectra.

[‡]Putative peptide homodimer formed by intermolecular disulfide bond of one of the two cysteines in this peptide.

[†]Intramolecular disulfide bond between vicinal cysteines.

*VC; vicinal cysteines Cys1669 or Cys1670. Disulfide can be formed to either Cys1669 or Cys1670 in the peptide, but the two possible products, while chemically distinct, are identical in mass.

A major concern was whether any non-literature disulfide pairs arose from internal rearrangements within a given A1A2A3 of a particular molecule before digestion or if they are the artifactual result of intermolecular crosslinking or intramolecular disulfide scrambling during the digestion and mass spectral assay. We generated a theoretical list of disulfide dimers, where the disulfide linkage is between two cysteines that are in the same peptide. Such an arrangement could only arise from an intermolecular crosslink, presumably a disulfide scrambling event during the procedure. In addition, this allows us to assess to some degree the odds of matching a predicted peak to the experimental data merely by chance. Accordingly, the predicted fragment ions listed in Table 3.4 were searched for in the MALDI mass spectral data using the Bruker software. As before, the predicted list was generated assuming that two missed cleavages were possible. Seventy calculated peptide pairs were used to screen an experimental list of 300 masses with a one Da error threshold. The detailed experimental mass list and the predicted cysteine peptide dimer list is given in the appendix 1. Three matches from the seventy calculated cysteine peptide pairs to the experimental mass list were detected in this particular experiment. In other experiments, an additional two matches were found. Note that there are multiple ways to generate a given disulfide dimer from peptides with various degrees of incomplete digestion, yet in each case we only observe a match for a particular arrangement. For example, there is an experimental mass within the error limit matching the calculated mass of 3133.46 for the disulfide linked mixed ‘dimer’ pair of CSQPLD / CSQPLDVILLLDGSSSFPASYFD. But there is no sign of either true dimer

CSQPLD / CSQPLD or of dimer CSQPLDVILLLDGSSSFPASYFD / CSQPLDVILLLDGSSSFPASYFD. This low number of matches between the calculated peptide mass list and the experimental mass list and the lack of families of dimers being found indicates that, if it is occurring at all, disulfide scrambling is very rare.

Table 3.4 Possible 'dimers' of GluC fragments which might be generated by disulfide scrambling of the A1A2A3 domains, assuming no more than two missed cleavages. Pairs with a match in the experimental data have their masses in bold.

Residue No.	Mass	Sequence	Residue No.	Mass	Sequence	Mass*
1270-1277	1016.49	FYCSRLLD	1270-1283	1716.90	FYCSRLLDLVFLLD	2770.39
1270-1277	1016.49	FYCSRLLD	1270-1277	1016.49	FYCSRLLD	2029.97
1270-1277	1016.49	FYCSRLLD	1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	3309.55
1270-1277	1016.49	FYCSRLLD	1270-1290	2433.25	FYCSRLLDLVFLLDGSSRLSE	3446.74
1270-1277	1016.49	FYCSRLLD	1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	3309.55
1664-1673	1107.53	LVLQRCCSGE	1661-1673	1390.65	APDLVLQRCCSGE	2495.17
1664-1673	1107.53	LVLQRCCSGE	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	3401.69
1664-1673	1107.53	LVLQRCCSGE	1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	4044.96
1686-1691	662.28	CSQPLD	1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	2511.20
1664-1673	1107.53	LVLQRCCSGE	1664-1673	1107.53	LVLQRCCSGE	2212.06
1686-1691	662.28	CSQPLD	1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	3599.71
1686-1691	662.28	CSQPLD	1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	***3133.46***
1686-1691	662.28	CSQPLD	1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	2511.20
1686-1691	662.28	CSQPLD	1686-1691	662.28	CSQPLD	1321.56
1453-1459	812.39	IVSYLCD	1453-1459	812.39	IVSYLCD	1621.77
1453-1459	812.39	IVSYLCD	1448-1459	1468.67	QQRDEIVSYLCD	2278.06
1453-1459	812.39	IVSYLCD	1452-1463	1351.65	EIVSYLCDLAPE	2161.03
1453-1459	812.39	IVSYLCD	1452-1459	941.43	EIVSYLCD	1750.81
1664-1673	1107.53	LVLQRCCSGE	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	3401.69
1452-1459	941.43	EIVSYLCD	1452-1459	941.43	EIVSYLCD	1879.86
1452-1459	941.43	EIVSYLCD	1453-1463	1222.60	IVSYLCDLAPE	2161.03
1452-1459	941.43	EIVSYLCD	1448-1459	1468.67	QQRDEIVSYLCD	2407.10
1452-1459	941.43	EIVSYLCD	1452-1463	1351.65	EIVSYLCDLAPE	***2290.07***

1452-1459	941.43	EIVSYLCD	1453-1459	812.39	IVSYLCD	1750.81
1453-1463	1222.60	IVSYLCDLAPE	1453-1459	812.39	IVSYLCD	2030.99
1453-1463	1222.60	IVSYLCDLAPE	1448-1459	1468.67	QQRDEIVSYLCD	2688.28
1453-1463	1222.60	IVSYLCDLAPE	1452-1459	941.43	EIVSYLCD	2161.03
1661-1673	1390.65	APDLVLQRCCSGE	1661-1673	1390.65	APDLVLQRCCSGE	2778.13
1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	3700.83
1453-1463	1222.60	IVSYLCDLAPE	1453-1463	1222.60	IVSYLCDLAPE	2442.20
1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	4591.33
1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	1664-1673	1107.53	LVLQRCCSGE	3401.69
1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	4367.26
1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	4429.20
1270-1283	1716.90	FYCSRLLDLVFLLD	1270-1283	1716.90	FYCSRLLDLVFLLD	3430.81
1270-1283	1716.90	FYCSRLLDLVFLLD	1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	4009.96
1270-1283	1716.90	FYCSRLLDLVFLLD	1270-1277	1016.49	FYCSRLLD	2730.39
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	4789.34
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	5877.85
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	5411.61
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	5484.77
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	5517.71
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1686-1697	1328.71	CSQPLDVILLLD	5234.59
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	5234.59
1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	1686-1691	662.28	CSQPLD	3599.71
1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	4891.36
1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	4768.37
1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	4323.10
1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	1686-1691	662.28	CSQPLD	3133.46
1686-1708	2474.18	CSQPLDVILLLDGSSSFPASYFD	1686-1697	1328.71	CSQPLDVILLLD	3799.89

1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	5033.69
1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	5095.63
1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	1686-1697	1328.71	CSQPLDVILLLD	3844.71
1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	4812.51
1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	1686-1691	662.28	CSQPLD	3177.63
1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	5157.56
1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	4874.44
1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	1674-1697	2518.35	GLQIPTLSPAPDCSQPLDVILLLD	5095.63
1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	1674-1691	1851.92	GLQIPTLSPAPDCSQPLD	4429.20
1656-1673	1986.97	TLPREAPDLVLQRCCSGE	1656-1673	1986.97	TLPREAPDLVLQRCCSGE	3970.95
1656-1673	1986.97	TLPREAPDLVLQRCCSGE	1664-1691	2940.43	LVLQRCCSGEGLQIPTLSPAPDCSQPLD	4924.40
1656-1673	1986.97	TLPREAPDLVLQRCCSGE	1661-1685	2580.28	APDLVLQRCCSGEGLQIPTLSPAPD	4564.25
1656-1673	1986.97	TLPREAPDLVLQRCCSGE	1664-1685	2297.16	LVLQRCCSGEGLQIPTLSPAPD	4281.14
1656-1673	1986.97	TLPREAPDLVLQRCCSGE	1661-1673	1390.65	APDLVLQRCCSGE	3376.62
1656-1673	1986.97	TLPREAPDLVLQRCCSGE	1664-1673	1107.53	LVLQRCCSGE	3091.50
1857-1881	2783.37	LPTMVTLGNSFLHKLCSGFVRICMD	1857-1881	2783.37	LPTMVTLGNSFLHKLCSGFVRICMD	5563.74
1857-1881	2783.37	LPTMVTLGNSFLHKLCSGFVRICMD	1857-1882	2912.49	LPTMVTLGNSFLHKLCSGFVRICMDE	5692.86
1857-1881	2783.37	LPTMVTLGNSFLHKLCSGFVRICMD	1856-1881	2898.46	DLPTMVTLGNSFLHKLCSGFVRICD	5681.83
1857-1881	2783.37	LPTMVTLGNSFLHKLCSGFVRICMD	1855-1881	3027.57	EDLPTMVTLGNSFLHKLCSGFVRICMD	5807.94
1857-1882	2912.49	LPTMVTLGNSFLHKLCSGFVRICMDE	1857-1882	2912.49	LPTMVTLGNSFLHKLCSGFVRICMDE	5821.97
1857-1882	2912.49	LPTMVTLGNSFLHKLCSGFVRICMDE	1855-1882	3156.67	EDLPTMVTLGNSFLHKLCSGFVRICMDE	6066.16
1857-1882	2912.49	LPTMVTLGNSFLHKLCSGFVRICMDE	1856-1881	2898.46	DLPTMVTLGNSFLHKLCSGFVRICD	5807.94
1856-1881	2898.46	DLPTMVTLGNSFLHKLCSGFVRICD	1856-1881	2898.46	DLPTMVTLGNSFLHKLCSGFVRICD	5793.92
1856-1881	2898.46	DLPTMVTLGNSFLHKLCSGFVRICD	1855-1881	3027.57	EDLPTMVTLGNSFLHKLCSGFVRICMD	5923.03
1856-1881	2898.46	DLPTMVTLGNSFLHKLCSGFVRICD	1855-1882	3156.67	EDLPTMVTLGNSFLHKLCSGFVRICMDE	6052.13
1855-1881	3027.57	EDLPTMVTLGNSFLHKLCSGFVRICMD	1855-1881	3027.57	EDLPTMVTLGNSFLHKLCSGFVRICMD	6051.14
1855-1881	3027.57	EDLPTMVTLGNSFLHKLCSGFVRICMD	1855-1882	3156.67	EDLPTMVTLGNSFLHKLCSGFVRICMD	6181.24

1855-1882	3156.67	EDLPTMVTLGNSFLHKLCSGFVRICMDE	1855-1882	3156.67	EDLPTMVTLGNSFLHKLCSGFVRICMDE	6310.37
1448-1459	1468.67	QQRDEIVSYLCD	1448-1459	1468.67	QQRDEIVSYLCD	2934.35
1448-1459	1468.67	QQRDEIVSYLCD	1452-1463	1351.65	EIVSYLCDLAPE	2817.32
1448-1459	1468.67	QQRDEIVSYLCD	1453-1463	1222.60	IVSYLCDLAPE	2688.28
1448-1459	1468.67	QQRDEIVSYLCD	1452-1459	941.43	EIVSYLCD	2407.10
1452-1463	1351.65	EIVSYLCDLAPE	1452-1463	1351.65	EIVSYLCDLAPE	2700.29
1452-1463	1351.65	EIVSYLCDLAPE	1448-1459	1468.67	QQRDEIVSYLCD	2161.03
1452-1463	1351.65	EIVSYLCDLAPE	1453-1463	1222.60	IVSYLCDLAPE	***2571.25***
1270- 1290	2433.25	FYCSRLLDLVFLLDGSSRLSE	1270- 1290	2433.25	FYCSRLLDLVFLLDGSSRLSE	4863.50
1270- 1290	2433.25	FYCSRLLDLVFLLDGSSRLSE	1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	4726.31
1270- 1290	2433.25	FYCSRLLDLVFLLDGSSRLSE	1270- 1283	1716.90	FYCSRLLDLVFLLD	4147.15
1270- 1290	2433.25	FYCSRLLDLVFLLDGSSRLSE	1270- 1277	1016.49	FYCSRLLD	3446.74
1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	4589.12
1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	1270- 1290	2433.25	FYCSRLLDLVFLLDGSSRLSE	4726.31
1260-1277	2296.06	EFDISEPPLHDFYCSRLLD	1270- 1277	1016.49	FYCSRLLD	3309.55

This data also allows us to establish the odds that other matches are false positives. If we assume that all of these were false hits (no scrambling) this would also suggest that, to a first approximation, we have a "random" hit rate of about 4-6% when searching hypothesis driven mass lists against our experimental data. With 75 candidate non-literature disulfide pairs listed in Table 3.3, we might expect 3-4 false positives.

According to the literature, Cys1272-Cys1458, Cys1686-1872, and Cys1669-Cys1670 are linked and would generate possible GluC fragments 1-13, 14-21, and 22-27 respectively in the Table 3.3. We wished to test the hypothesis that VWF could have alternative disulfide rearrangements, listed as entries 28-102 in the Table 3.3. We conducted experiments with reduced/oxidized glutathione ratios below and above the levels found in plasma and also under fully oxidized conditions without any glutathione in any form present. Partially oxidized conditions were examined using a reduced/oxidized glutathione ratio of 1/1000. Reduced conditions were represented using reduced/oxidized glutathione of ratio 1/1.

All the masses which were observed that correspond to a theoretical disulfide are summarized in Table 3.5 for the digest in the absence of glutathione, in Table 3.6 for GSH/GSSG 1/1, and Table 3.7 for GSH/GSSG 1/1000. The MALDI-TOF-MS, Figure 3.7, of the VWF GluC digested in the absence of either oxidized or reduced glutathione was found to contain fragments with masses the same as those predicted for two disulfide linked peptides containing Cys1272-Cys1458. The fact that two versions of this disulfide linkage are found differing by a missed cleavage at an aspartate raises confidence that this is a real signal.

However no fragments with a mass appropriate for a disulfide linked peptide containing Cys1686-Cys1872 or Cys1669-Cys1670 were observed. This is not evidence that these fragments are missing because the disulfides are completely absent. There are many peptides which ionize poorly enough that they are not found in the MS data. On the other hand, it is interesting to note that in some of the digests a fragment with a mass of approximately 1986 Da was found. This corresponds to the mass expected of the predicted peptide 1656-1673, TLPREAPDLVLQRCCSGE, which contains Cys1669 and Cys1670. But at this mass the two cysteines are in reduced form. Similarly, in some of the digests there are fragments with masses which correspond to those expected for predicted peptides contain each of these cysteines in reduced form.

As seen in Table 3.6, in the presence of 1/1 GSH/GSSG there are also two masses which correspond to those predicted for a 1272-1458 disulfide. One of these is the same as in Table 3.5 and one is different. Interestingly, there is a single peak which corresponds to a peptide with 1669-1670 in the oxidized form and a single peak that has a mass for a peptide pair corresponding to the third literature disulfide between 1686 and 1872. However, as summarized in Table 3.7, in the presence of glutathione in an reduced to oxidized ratio of 1/1000, we failed to detect any of the disulfide containing peptides predicted in the literature.

More interestingly, there is clear evidence for previously unreported disulfides in the data. Under all three conditions we find one or more fragments with masses that are consistent with peptides linked by a disulfide between cysteine 1458 and one of the vicinal cysteines at 1669 or 1670. Obviously the mass is identical regardless of whether

1669 or 1670 makes the bond. In the presence of 1/1 GSH/GSSG or in the absence of any oxidized or reduced glutathione, we see masses that are consistent with the presence of peptides linked between 1272 and either 1669 or 1670. It is noteworthy that these linkages seemed to be confirmed by multiple digest products, increasing confidence that they are true hits and not false positives.

Table 3.5. The observed fragments corresponding to different disulfide connectivity in A1A2A3 domain after GluC digestion in the absence of glutathione. Disulfides that are not found in the literature are in **bold**.

No Glutathione							
Disulfide Position	**Intensity**	**m/z+**	**m/z**	**ppm Error**	**Missed cleavage**	**Domain**	**Sequence**
Cys1272-Cys1458	0.88	3242.62	3242.61	3.70	2	A1	FYCSRLLDLVFLLDGSSPLSE / VSYLCD
	0.63	3250.57	3250.54	9.23	2	A1	FYCSRLLD / EIVSYLCDLAPEAPPPTLPPD
Cys1272-VC	0.53	3008.53	3008.42	37.23	1	A1A2	ISEPPLHDFYCSRLLD / LVLQRCCSGE
	0.33	3536.78	3536.75	8.48	2	A1A2	FYCSRLLDLVFLLDGSSRLSE / LVLQRCCSGE
Cys1458-VC	12.74	2571.22	2571.17	19.06	2	A1A2	LVLQRCCSGE / QQRDEIVSYLCD
	0.93	2044.93	2044.51	205.39	1	A1A2	LVLQRCCSGE / EIVSYLCD
	0.56	3387.86	3388.64	-228.46	2	A1A2	APDLVLQRCCSGEGLQIPTLSPAPD / VSYLCD
	0.27	3746.77	3746.77	2.14	2	A1A2	LVLQRCCSGEGLQIPTLSPAPDCSQPLDVILLLD / VSYLCD

Table 3.6. The observed fragments corresponding to different disulfide connectivity in A1A2A3 domain after GluC digestion in the presence of glutathione in a reduced/oxidized ratio of 1/1. Disulfides that are not found in the literature are in **bold**.

Glutathione Reduced/Oxidized 1/1							
Disulfide Position	**Intensity**	**m/z+**	**m/z**	**ppm Error**	**Missed Cleavage**	**Domain**	**Sequence**
Cys1272-Cys1458	0.13	3241.60	3241.13	146.53	2	A1	FYCSRLLDLVFLLDGSSRLSE/ VSYLCD
	0.23	2714.39	2714.28	42.74	1	A1	ISEPPLHDFYCSRLL / VSYLCD
Cys1669-Cys1670	0.56	1985.96	1985.19	389.23	2	A2	TLPREAPDVLVLQRCCSGE
Cys1686-Cys1872	0.18	3007.44	3007.40	8.97	0	A3	CSQPLD/LPTMVTLGNSFLHKLCSFGVRICMD
Cys1272-VC	0.18	3008.44	3008.41	6.64	1	A1A2	ISEPPLHDFYCSRLL / VLQRCCSGE
Cys1458-VC	3.68	2571.33	2571.16	63.39	2	A1A2	LVLQRCCSGE / QQRDEIVSYLCD
	0.17	2044.93	2044.77	78.24	1	A1A2	LVLQRCCSGE / EIVSYLCD

Table 3.7. The observed fragments corresponding to different disulfide connectivity in A1A2A3 domain after GluC digestion in the presence of glutathione in a reduced/oxidized ratio of 1/1000. Disulfides that are not found in the literature are in **bold**.

Glutathione Reduced/Oxidized 1/1000							
Disulfide Position	**Intensity**	**m/z+**	**m/z**	**ppm error**	**Missed Cleavage**	**Domain**	**Sequence**
Cys1458-VC	5.07	2571.39	2571.17	86.33	2	A1A2	QQRDEIVSYLD / LVLQRCCSGE

The consistent observation of the breaking of the vicinal disulfide bond in A2 domain between positions 1669-1670 and the formation of a new disulfide bond with one of the two cysteine residues of A1 domain. The cysteine residue in the 1458 position of A1 domain is more proximal to the vicinal cysteine residues in 1669-1670 positions, at least in primary structure, than the cysteine residue in the position 1272 of the A1 domain. The quaternary structure of these domains is not clear. In our experiments we observed higher intensities of the cysteine residue 1458 than the more distal cysteine residue located in the 1272 position, but the MS intensities are influenced by many factors and cannot be relied on for actual amounts of the two species that are present.

There is no clear evidence that the redox potential makes these alternative arrangements more or less likely. There was significant variations in what we found, but no real trend linked to glutathione levels. Rather, the results seem more dependent on the quality of the particular experiment conducted. Nevertheless, our results are consistent with a redox switch in 1669-1670 position which is possibly triggered by shear stress or by fluctuations of the plasma thiol /disulfide pools.

CONCLUSIONS

Over the past few years researchers have just begun to understand the redox dependent regulation of the disulfide bonds in protein activity. The regulation of VWF structure and function is important in the context of cardiovascular diseases since they are associated with elevated VWF activity. Improvement of our understanding of the structure, function and conformation of VWF could possibly make VWF a drug target, so advancement in this area of an interest to the pharmaceutical industry. In conclusion

the disulfide bond structure of proteins and the redox dependence of the disulfide sturucture can be mapped using the method described in this study. Our results indicated the presence of disulfide arrangements within the A1A2A3 domains of VWF differing from those found in the crystal structures. The consistent observation of the breaking of the vicinal disulfide bond in A2 domain between positions 1669-1670 and the formation of a new disulfide bond with one of the two cysteine residues of A1 domain suggests a redox switch in the 1669-1670 position. The results presented here demonstrate the need for further research into redox activated switches in plasma proteins. In this study we used VWF isolated from the human blood. In the future experiments VWF isolated from other species might be used and compared with the experimental findings of this study. Experiments carried under shear stress using a broader GSH/GSSG concentrations need to be considered. Redox dependence of the disulfide bonding pattern of VWF could be tested using another plasma thiol/disulfide couple, for example cysteine/cysteine in future experiments.

REFERENCES

1. Kadokura, H., Katzen, F., and Beckwith, J. (2003) Protein disulfide bond formation in prokaryotes, *Annual Review of Biochemistry 72*, 111-135.

2. Freedman, R. B. (1995) THE FORMATION OF PROTEIN DISULFIDE BONDS, *Current Opinion in Structural Biology 5*, 85-91.

3. Anfinsen, C. B., Haber, E., Sela, M., and White, F. H. (1961) KINETICS OF FORMATION OF NATIVE RIBONUCLEASE DURING OXIDATION OF REDUCED POLYPEPTIDE CHAIN, *Proceedings of the National Academy of Sciences of the United States of America 47*, 1309.

4. Berndt, C., Lillig, C. H., and Holmgren, A. (2008) Thioredoxins and glutaredoxins as facilitators of protein folding, *Biochimica Et Biophysica Acta-Molecular Cell Research 1783*, 641-650.

5. Rudolph, R., and Lilie, H. (1996) In vitro folding of inclusion body proteins, *Faseb Journal 10*, 49-56.

6. Hogg, P. J. (2003) Disulfide bonds as switches for protein function, *Trends in Biochemical Sciences 28*, 210-214.

7. Luken, B. M. (2008) Extracellular control of VWF multimer size and thiol-disulfide exchange, *Journal of Thrombosis and Haemostasis 6*, 1131-1134.

8. Schmidt, B., Ho, L., and Hogg, P. J. (2006) Allosteric disulfide bonds, *Biochemistry 45*, 7429-7433.

9. Essex, D. W., and Li, M. R. (2003) Redox control of platelet aggregation, *Biochemistry 42*, 129-136.

10. Ahamed, J., Versteeg, H. H., Kerver, M., Chen, V. M., Mueller, B. M., Hogg, P. J., and Ruf, W. (2006) Disulfide isomerization switches tissue factor from coagulation to cell signaling, *Proceedings of the National Academy of Sciences of the United States of America 103*, 13932-13937.

11. Pendurthi, U., Ghosh, S., Mandal, S., and Rao, L. V. M. (2006) Tissue factor activation: Is disulfide switching a general regulatory mechanism?, *Blood 108*, 1747.

12. Bach, R. R. (2006) Tissue factor encryption, *Arteriosclerosis Thrombosis and Vascular Biology 26*, 456-461.

13. Stathakis, P., Lay, A. J., Fitzgerald, M., Schlieker, C., Matthias, L. J., and Hogg, P. J. (1999) Angiostatin formation involves disulfide bond reduction and

proteolysis in kringle 5 of plasmin, *Journal of Biological Chemistry 274*, 8910-8916.

14. Carugo, O., Cemazar, M., Zahariev, S., Hudaky, I., Gaspari, Z., Perczel, A., and Pongor, S. (2003) Vicinal disulfide turns, *Protein Engineering 16*, 637-639.

15. Zhang, Q., Zhou, Y. F., Zhang, C. Z., Zhang, X. H., Lu, C. F., and Springer, T. A. (2009) Structural specializations of A2, a force-sensing domain in the ultralarge vascular protein von Willebrand factor, *Proceedings of the National Academy of Sciences of the United States of America 106*, 9226-9231.

16. Ruggles, E. L., and Hondal, R. J. (2006) Synthesis and properties of disulfide-bond containing eight-membered rings, *Tetrahedron Letters 47*, 4281-4284.

17. Wang, X. H., Connor, M., Smith, R., Maciejewski, M. W., Howden, M. E. H., Nicholson, G. M., Christie, M. J., and King, G. F. (2000) Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge, *Nature Structural Biology 7*, 505-513.

18. Blake, C. C. F., Ghosh, M., Harlos, K., Avezoux, A., and Anthony, C. (1994) THE ACTIVE-SITE OF METHANOL DEHYDROGENASE CONTAINS A DISULFIDE BRIDGE BETWEEN ADJACENT CYSTEINE RESIDUES, *Nature Structural Biology 1*, 102-105.

19. Stamnaes, J., Pinkas, D. M., Fleckenstein, B., Khosla, C., and Sollid, L. M. (2010) Redox Regulation of Transglutaminase 2 Activity, *Journal of Biological Chemistry 285*, 25402-25409.

20. Passam FH, R. S., Qi M, Raftery MJ, Wong JW, Tanaka K, Ioannou Y, Zhang JY, Gemmell R, Qi JC, Giannakopoulos B, Hughes WE, Hogg PJ, Krilis SA. (2010) Redox control of β2-glycoprotein I-von Willebrand factor interaction by thioredoxin-1., *J Thromb Haemost 8*, 1754-1762.

21. Choi, H., Aboulfatova, K., Pownall, H. J., Cook, R., and Dong, J. F. (2007) Shear-induced disulfide bond formation regulates adhesion activity of von Willebrand factor, *Journal of Biological Chemistry 282*, 35604-35611.

22. Yeh, H. C., Zhou, Z., Choi, H., Tekeoglu, S., May, W., Wang, C., Turner, N., Scheiflinger, F., Moake, J. L., and Dong, J. F. (2010) Disulfide bond reduction of von Willebrand factor by ADAMTS-13, *Journal of Thrombosis and Haemostasis 8*, 2778-2788.

23. Xie, L. J., Chesterman, C. N., and Hogg, P. J. (2000) Reduction of von Willebrand factor by endothelial cells, *Thrombosis and Haemostasis 84*, 506-513.

24. Xie, L. J., Chesterman, C. N., and Hogg, P. J. (2001) Control of von Willebrand factor multimer size by thrombospondin-1, *Journal of Experimental Medicine 193*, 1341-1349.

25. Ferrari, D. M., and Soling, H. D. (1999) The protein disulphide-isomerase family: unravelling a string of folds, *Biochemical Journal 339*, 1-10.

26. Holmgren, A., Johansson, C., Berndt, C., Lonn, M. E., Hudemann, C., and Lillig, C. H. (2005) Thiol redox control via thioredoxin and glutaredoxin systems, *Biochemical Society Transactions 33*, 1375-1377.

27. Brosnan, J. T., and Brosnan, M. E. (2006) The sulfur-containing amino acids: An overview, *Journal of Nutrition 136*, 1636S-1640S.

28. Jones, D. P. (2002) Redox potential of GSH/GSSG couple: Assay and biological significance, *Enzymology 348*, 93-112.

29. White, A. C., Thannickal, V. J., and Fanburg, B. L. (1994) GLUTATHIONE DEFICIENCY IN HUMAN-DISEASE, *Journal of Nutritional Biochemistry 5*, 218-226.

30. Flagg, E. W., Coates, R. J., Jones, D. P., Eley, J. W., Gunter, E. W., Jackson, B., and Greenberg, R. S. (1993) PLASMA TOTAL GLUTATHIONE IN HUMANS AND ITS ASSOCIATION WITH DEMOGRAPHIC AND HEALTH-RELATED FACTORS, *British Journal of Nutrition 70*, 797-808.

31. Garner, M., Reglinski, J., Smith, W. E., and McMurray, J. (1992) OXIDATION-STATE OF GLUTATHIONE IN THE ERYTHROCYTE, *Clinical Science 83*, 637-637.

32. Smith, J. E. (1974) RELATIONSHIP OF IN-VIVO ERYTHROCYTE GLUTATHIONE FLUX TO OXIDIZED GLUTATHIONE TRANSPORT-SYSTEM, *Journal of Laboratory and Clinical Medicine 83*, 444-450.

33. Hagenfeldt, L., Arvidsson, A., and Larsson, A. (1978) GLUTATHIONE AND GAMMA-GLUTAMYLCYSTEINE IN WHOLE-BLOOD, PLASMA AND ERYTHROCYTES, *Clinica Chimica Acta 85*, 167-173.

34. Richie, J. P., Skowronski, L., Abraham, P., and Leutzinger, Y. (1996) Blood glutathione concentrations in a large-scale human study, *Clinical Chemistry 42*, 64-70.

35. Jones, D. P., Go, Y. M., Anderson, C. L., Ziegler, T. R., Kinkade, J. M., and Kirlin, W. G. (2004) Cysteine/cystine couple is a newly recognized node in the circuitry for biologic redox signaling and control, *Faseb Journal 18*, 1246.

36. Jones, D. P., Carlson, J. L., Samiec, P. S., Sternberg, P., Mody, V. C., Reed, R. L., and Brown, L. A. S. (1998) Glutathione measurement in human plasma Evaluation of sample collection, storage and derivatization conditions for analysis of dansyl derivatives by HPLC, *Clinica Chimica Acta 275*, 175-184.

37. Svardal, A. M., Mansoor, M. A., and Ueland, P. M. (1990) DETERMINATION OF REDUCED, OXIDIZED, AND PROTEIN-BOUND GLUTATHIONE IN HUMAN PLASMA WITH PRECOLUMN DERIVATIZATION WITH MONOBROMOBIMANE AND LIQUID-CHROMATOGRAPHY, *Analytical Biochemistry 184*, 338-346.

38. Beutler, E., and Gelbart, T. (1985) PLASMA GLUTATHIONE IN HEALTH AND IN PATIENTS WITH MALIGNANT DISEASE, *Journal of Laboratory and Clinical Medicine 105*, 581-584.

39. Griffith, O. W., and Meister, A. (1979) GLUTATHIONE - INTERORGAN TRANSLOCATION, TURNOVER, AND METABOLISM, *Proceedings of the National Academy of Sciences of the United States of America 76*, 5606-5610.

40. Bayele, H. K., Murdock, P. J., Perry, D. J., and Pasi, K. J. (2002) Simple shifts in redox/thiol balance that perturb blood coagulation, *Febs Letters 510*, 67-70.

41. Hagen, T. M., Wierzbicka, G. T., Sillau, A. H., Bowman, B. B., and Jones, D. P. (1990) BIOAVAILABILITY OF DIETARY GLUTATHIONE - EFFECT ON PLASMA-CONCENTRATION, *American Journal of Physiology 259*, G524-G529.

42. Chawla, R. K., Lewis, F. W., Kutner, M. H., Bate, D. M., Roy, R. G. B., and Rudman, D. (1984) PLASMA CYSTEINE, CYSTINE, AND GLUTATHIONE IN CIRRHOSIS, *Gastroenterology 87*, 770-776.

43. Mills, B. J., Weiss, M. M., Lang, C. A., Liu, M. C., and Ziegler, C. (2000) Blood glutathione and cysteine changes in cardiovascular disease, *Journal of Laboratory and Clinical Medicine 135*, 396-401.

44. Huizinga, E. G., Tsuji, S., Romijn, R. A. P., Schiphorst, M. E., de Groot, P. G., Sixma, J. J., and Gros, P. (2002) Structures of glycoprotein Ib alpha and its complex with von Willebrand factor A1 domain, *Science 297*, 1176-1179.

45. Huizinga, E. G., Martijn van der Plas, R., Kroon, J., Sixma, J.J., Gros, P. (1997) Crystal structure of the A3 domain of human von Willebrand factor: implications for collagen binding., *Structure 5*, 1147-1156.

46. Luken, B. M., Winn, L. Y. N., Emsley, J., Lane, D. A., and Crawley, J. T. B. (2010) The importance of vicinal cysteines, C1669 and C1670, for von Willebrand factor A2 domain function, *Blood 115*, 4910-4913.

47. Mauri, P., Toppo, S., De Palma, A., Benazzi, L., Maiorino, M., and Ursini, F. (2010) IDENTIFICATION BY MS/MS OF DISULFIDES PRODUCED BY A FUNCTIONAL REDOX TRANSITION, in *Methods in Enzymology, Vol 473: Thiol Redox Transitions in Cell Signaling, Pt a: Chemistry and Biochemistry of Low Molecular Weight and Protein Thiols*, pp 217-225, Elsevier Academic Press Inc, San Diego.

48. Tang, H. Y., and Speicher, D. W. (2004) Determination of disulfide-bond linkages in proteins, *Curr Protoc Protein Sci Chapter 11*, Unit 1111.

49. Chang, J. Y., Lu, B. Y., and Li, L. (2005) Conformational impurity of disulfide proteins: detection, quantification, and properties, *Anal Biochem 342*, 78-85.

50. Moulaei, T., Stuchlik, O., Reed, M., Yuan, W., Pohl, J., Lu, W., Haugh-Krumpe, L., O'Keefe, B. R., and Wlodawer, A. (2010) Topology of the disulfide bonds in the antiviral lectin scytovirin, *Protein Sci 19*, 1649-1661.

51. Sorensen, S. B., Sorensen, T. L., and Breddam, K. (1991) FRAGMENTATION OF PROTEINS BY STAPHYLOCOCCUS-AUREUS STRAIN-V8 PROTEASE - AMMONIUM BICARBONATE STRONGLY INHIBITS THE ENZYME BUT DOES NOT IMPROVE THE SELECTIVITY FOR GLUTAMIC-ACID, *Febs Letters 294*, 195-197.

APPENDIX

Supplemental Table 1. Experimental mass list for Glu C digest (1:1 GSH:GSSG). Highlighted peaks are those which match a predicted mass in Table 3.4.

m/z	Intensity	Relative Intensity	Calculated Peptide Masses
976.40	6579.00	0.01	1321.56
1014.36	40933.00	0.07	1621.77
1030.35	16077.19	0.03	1750.81
1036.36	7348.50	0.01	1750.81
1052.37	15127.89	0.02	1879.86
1069.42	85970.79	0.14	2029.97
1073.37	2332.81	0.00	2030.99
1076.50	3366.15	0.01	2161.03
1081.65	2762.29	0.00	2161.03
1091.43	20769.56	0.03	2161.03
1107.41	7135.91	0.01	2212.06
1117.60	74445.11	0.12	2290.07
1123.78	1068.08	0.00	2407.10
1134.53	4584.70	0.01	2407.10
1139.59	29707.10	0.05	2442.20
1147.57	3145.45	0.01	2495.17
1155.56	8709.34	0.01	2511.20
1161.60	11283.61	0.02	2571.25
1177.56	3301.29	0.01	2688.28
1205.61	4125.85	0.01	2700.29
1207.49	2959.76	0.00	2730.39
1223.68	30452.64	0.05	2770.39
1226.44	1944.82	0.00	2778.13
1230.56	11885.04	0.02	2817.32
1234.56	1744.65	0.00	2984.35
1235.59	2386.76	0.00	3091.50
1244.68	3582.24	0.01	3133.46
1245.65	10469.04	0.02	3177.63
1248.56	1540.70	0.00	3309.55
1260.67	38274.06	0.06	3309.55
1264.09	3171.25	0.01	3309.55
1281.73	17778.00	0.03	3401.69
1282.62	13089.00	0.02	3401.69
1289.66	5522.96	0.01	3430.81
1298.61	4978.01	0.01	3446.74
1313.64	7844.00	0.01	3446.74

1319.65	10565.82	0.02	3599.71
1321.55	12903.55	0.02	3700.83
1328.72	9835.00	0.02	3799.89
1343.57	4368.88	0.01	3844.71
1356.72	31494.42	0.05	3970.95
1358.87	7612.73	0.01	4009.96
1359.78	47987.79	0.08	4147.15
1374.72	619404.12	1.00	4281.14
1376.74	104342.12	0.17	4323.10
1377.78	21400.02	0.03	4367.26
1386.77	130317.95	0.21	4429.20
1396.74	61350.17	0.10	4564.25
1402.75	8845.46	0.01	4589.12
1412.71	20372.84	0.03	4591.33
1417.75	17778.00	0.03	4726.31
1435.69	6940.29	0.01	4726.31
1450.68	46463.62	0.08	4768.37
1466.68	134615.76	0.22	4789.34
1488.70	38294.02	0.06	4812.51
1504.66	11073.56	0.02	4863.50
1510.76	34036.87	0.05	4874.44
1518.82	110082.13	0.18	4891.36
1529.76	5693.19	0.01	4924.40
1534.79	41450.00	0.07	5033.69
1538.79	70456.58	0.11	5095.63
1551.74	4551.63	0.01	5095.63
1566.75	8546.40	0.01	5157.56
1583.30	2333.64	0.00	5234.59
1594.82	7858.22	0.01	5234.59
1597.69	27425.33	0.04	5411.61
1627.78	3120.39	0.01	5484.77
1646.76	8441.88	0.01	5517.71
1660.73	10009.23	0.02	5877.85
1665.71	15753.39	0.03	
1675.77	7879.37	0.01	
1692.82	163269.28	0.26	
1714.81	26994.54	0.04	
1730.04	10273.04	0.02	
1736.82	2813.98	0.00	
1745.84	30603.59	0.05	
1762.86	9880.00	0.02	
1768.90	4863.69	0.01	
1775.85	10739.00	0.02	
1792.84	6308.45	0.01	

1804.87	7722.83	0.01
1810.89	1628.17	0.00
1827.28	4661.92	0.01
1837.89	62294.74	0.10
1848.96	10499.00	0.02
1859.88	11977.00	0.02
1864.92	16571.89	0.03
1884.56	2165.04	0.00
1908.00	16771.00	0.03
1924.03	241242.18	0.39
1946.02	26325.78	0.04
1961.98	6915.44	0.01
1972.91	9992.11	0.02
1978.97	16755.34	0.03
1990.00	146487.19	0.24
1999.99	15255.00	0.02
2012.01	26519.00	0.04
2023.00	19681.00	0.03
2044.51	5771.25	0.01
2056.98	4082.77	0.01
2078.03	2600.41	0.00
2081.14	4421.00	0.01
2086.32	3015.78	0.00
2096.09	197143.92	0.32
2118.06	15606.32	0.03
2134.03	5033.03	0.01
2140.79	5693.38	0.01
2179.10	16817.00	0.03
2197.10	12063.00	0.02
2234.12	2926.96	0.00
2250.10	91004.84	0.15
2253.15	92029.42	0.15
2268.18	586200.16	0.95
2272.19	127853.33	0.21
2282.16	33936.50	0.05
2285.17	***26757.86***	***0.04***
2290.14	194800.10	0.31
2295.12	9070.99	0.01
2300.17	66366.28	0.11
2306.11	87697.06	0.14
2312.12	30167.27	0.05
2314.16	8736.82	0.01
2320.14	4138.82	0.01
2322.12	15656.00	0.03

2328.09	12549.57	0.02
2332.25	6948.43	0.01
2345.09	17603.00	0.03
2354.13	4253.72	0.01
2376.19	10236.00	0.02
2404.15	10128.82	0.02
2422.15	11588.24	0.02
2426.18	21289.20	0.03
2429.25	4538.47	0.01
2443.20	135767.60	0.22
2446.20	38194.28	0.06
2465.13	2326.86	0.00
2479.20	2451.49	0.00
2505.27	4400.28	0.01
2509.37	7638.94	0.01
2516.36	2682.93	0.00
2531.24	30814.83	0.05
2534.25	27650.76	0.04
2549.29	421126.78	0.68
2554.28	29298.86	0.05
2561.21	5219.44	0.01
2565.26	***170473.60***	***0.28***
2571.22	78890.59	0.13
2575.19	13249.66	0.02
2581.27	9223.94	0.01
2587.20	41033.42	0.07
2593.23	6963.07	0.01
2597.25	12876.55	0.02
2604.28	9760.94	0.02
2629.29	8827.00	0.01
2639.49	23727.13	0.04
2655.37	8734.00	0.01
2673.33	8989.00	0.01
2685.29	6964.80	0.01
2701.31	7797.00	0.01
2702.31	8949.00	0.01
2727.42	7097.00	0.01
2734.37	2395.28	0.00
2738.32	6430.68	0.01
2753.32	1714.70	0.00
2796.46	1770.92	0.00
2807.38	4594.00	0.01
2821.41	548.09	0.00
2838.39	995.32	0.00

2842.49	2057.03	0.00
2850.42	3696.00	0.01
2866.31	595.62	0.00
2900.36	805.09	0.00
2919.49	1661.18	0.00
2980.53	2107.56	0.00
2996.48	1356.24	0.00
3002.58	1299.29	0.00
3008.53	3311.30	0.01
3023.47	10154.01	0.02
3040.46	124890.61	0.20
3062.39	5188.84	0.01
3078.40	1977.70	0.00
3094.52	1040.30	0.00
3110.65	3451.28	0.01
3152.63	4700.00	0.01
3176.56	1706.77	0.00
3191.65	10092.68	0.02
3194.71	2623.96	0.00
3202.62	8355.00	0.01
3208.66	94636.81	0.15
3212.57	7683.62	0.01
3226.32	1414.54	0.00
3232.57	932.68	0.00
3242.62	5450.00	0.01
3250.54	3883.00	0.01
3265.56	3443.00	0.01
3344.75	3056.00	0.00
3360.72	2765.00	0.00
3377.31	2120.74	0.00
3387.86	3439.00	0.01
3404.67	2668.00	0.00
3416.74	2273.00	0.00
3426.72	4294.00	0.01
3457.78	2103.00	0.00
3473.82	2363.65	0.00
3496.71	2015.00	0.00
3508.73	1769.00	0.00
3522.78	1857.00	0.00
3536.78	2033.00	0.00
3551.62	1688.00	0.00
3552.67	1737.00	0.00
3572.71	1735.00	0.00
3582.89	1494.00	0.00

3600.80	1628.00	0.00
3607.94	1948.00	0.00
3623.97	550.01	0.00
3634.95	1798.00	0.00
3642.13	531.77	0.00
3652.85	1826.00	0.00
3666.96	1801.00	0.00
3680.96	1678.00	0.00
3698.92	2264.99	0.00
3715.98	2052.00	0.00
3735.91	2064.00	0.00
3746.81	1694.00	0.00
3755.03	1630.00	0.00
3780.56	2112.00	0.00
3791.02	1520.97	0.00
3796.04	571.81	0.00
3822.96	***533.16***	***0.00***
3845.00	1739.00	0.00
3860.22	553.22	0.00
3877.77	1928.64	0.00
3923.19	482.56	0.00
3943.28	758.71	0.00
3966.14	794.63	0.00
3968.17	3009.88	0.00
3979.13	994.25	0.00
3994.30	1948.00	0.00
4021.26	2322.00	0.00
4043.20	2100.00	0.00
4056.18	2068.00	0.00
4111.22	776.02	0.00
4126.32	563.45	0.00
4144.16	478.41	0.00
4162.27	29110.00	0.05
4184.21	2644.00	0.00
4206.03	2580.00	0.00
4273.34	1310.00	0.00
4289.56	1108.00	0.00
4297.44	1405.00	0.00
4313.37	1397.00	0.00
4340.54	1614.00	0.00
4350.45	1107.00	0.00
4365.62	1364.00	0.00
4408.46	1648.00	0.00
4414.49	2017.00	0.00

4424.58	633.38	0.00
4449.48	1081.00	0.00
4465.51	1279.00	0.00
4493.53	965.00	0.00
4535.72	553.55	0.00
4595.57	1794.00	0.00
4610.81	776.26	0.00
4627.66	1131.00	0.00
4644.37	1091.00	0.00
4692.81	541.36	0.00
4708.88	5853.44	0.01
4790.97	1727.00	0.00
4817.00	635.74	0.00
4833.75	1486.00	0.00
4847.95	1393.00	0.00
4856.38	377.93	0.00
4873.77	1161.00	0.00
4887.94	1358.00	0.00
4901.18	866.00	0.00
4918.35	1043.00	0.00
4931.83	1024.00	0.00
4971.23	969.00	0.00
4992.94	964.00	0.00
5029.92	909.00	0.00
5033.34	1540.00	0.00
5065.05	1229.00	0.00
5098.32	168.19	0.00
5113.53	999.00	0.00
5137.23	1292.00	0.00
5151.23	1274.00	0.00
5176.26	1388.00	0.00
5231.16	1030.00	0.00
5282.30	1001.00	0.00
5291.59	820.00	0.00
5309.09	161.76	0.00
5361.43	815.00	0.00
5377.91	941.00	0.00
5443.11	598.00	0.00
5477.38	851.00	0.00
5507.06	891.00	0.00
5590.52	990.00	0.00
5647.61	2365.00	0.00
5661.93	1907.72	0.00
5759.33	969.00	0.00

5791.66	832.00	0.00
5798.52	797.00	0.00
5813.75	939.00	0.00
5830.14	699.11	0.00
5892.03	969.00	0.00
5966.57	688.00	0.00
6008.12	806.00	0.00
6007.01	685.00	0.00
6158.56	604.00	0.00
6688.05	563.00	0.00
6698.72	578.00	0.00
6775.68	551.00	0.00
6778.28	522.00	0.00
6951.44	537.00	0.00
7089.48	541.00	0.00
7136.19	561.00	0.00
7151.38	816.00	0.00
7339.90	561.00	0.00
7342.40	632.00	0.00
7544.56	660.00	0.00
7559.28	94.45	0.00
7659.44	564.00	0.00
7804.92	733.00	0.00
7809.19	962.00	0.00
8091.21	576.00	0.00
8094.03	614.00	0.00

m/z	Intensity	Relative Intensity	Calculated Peptide Masses
976.40	6579.00	0.01	1321.56
1014.36	40933.00	0.07	1621.77
1030.35	16077.19	0.03	1750.81
1036.36	7348.50	0.01	1750.81
1052.37	15127.89	0.02	1879.86
1069.42	85970.79	0.14	2029.97
1073.37	2332.81	0.00	2030.99
1076.50	3366.15	0.01	2161.03
1081.65	2762.29	0.00	2161.03
1091.43	20769.56	0.03	2161.03
1107.41	7135.91	0.01	2212.06
1117.60	74445.11	0.12	2290.07
1123.78	1068.08	0.00	2407.10
1134.53	4584.70	0.01	2407.10
1139.59	29707.10	0.05	2442.20
1147.57	3145.45	0.01	2495.17
1155.56	8709.34	0.01	2511.20
1161.60	11283.61	0.02	2571.25
1177.56	3301.29	0.01	2688.28
1205.61	4125.85	0.01	2700.29
1207.49	2959.76	0.00	2730.39
1223.68	30452.64	0.05	2770.39
1226.44	1944.82	0.00	2778.13
1230.56	11885.04	0.02	2817.32
1234.56	1744.65	0.00	2984.35
1235.59	2386.76	0.00	3091.50
1244.68	3582.24	0.01	3133.46
1245.65	10469.04	0.02	3177.63
1248.56	1540.70	0.00	3309.55
1260.67	38274.06	0.06	3309.55
1264.09	3171.25	0.01	3309.55
1281.73	17778.00	0.03	3401.69
1282.62	13089.00	0.02	3401.69
1289.66	5522.96	0.01	3430.81
1298.61	4978.01	0.01	3446.74
1313.64	7844.00	0.01	3446.74
1319.65	10565.82	0.02	3599.71
1321.55	12903.55	0.02	3700.83
1328.72	9835.00	0.02	3799.89
1343.57	4368.88	0.01	3844.71
1356.72	31494.42	0.05	3970.95
1358.87	7612.73	0.01	4009.96
1359.78	47987.79	0.08	4147.15

1374.72	619404.12	1.00	4281.14
1376.74	104342.12	0.17	4323.10
1377.78	21400.02	0.03	4367.26
1386.77	130317.95	0.21	4429.20
1396.74	61350.17	0.10	4564.25
1402.75	8845.46	0.01	4589.12
1412.71	20372.84	0.03	4591.33
1417.75	17778.00	0.03	4726.31
1435.69	6940.29	0.01	4726.31
1450.68	46463.62	0.08	4768.37
1466.68	134615.76	0.22	4789.34
1488.70	38294.02	0.06	4812.51
1504.66	11073.56	0.02	4863.50
1510.76	34036.87	0.05	4874.44
1518.82	110082.13	0.18	4891.36
1529.76	5693.19	0.01	4924.40
1534.79	41450.00	0.07	5033.69
1538.79	70456.58	0.11	5095.63
1551.74	4551.63	0.01	5095.63
1566.75	8546.40	0.01	5157.56
1583.30	2333.64	0.00	5234.59
1594.82	7858.22	0.01	5234.59
1597.69	27425.33	0.04	5411.61
1627.78	3120.39	0.01	5484.77
1646.76	8441.88	0.01	5517.71
1660.73	10009.23	0.02	5877.85
1665.71	15753.39	0.03	
1675.77	7879.37	0.01	
1692.82	163269.28	0.26	
1714.81	26994.54	0.04	
1730.04	10273.04	0.02	
1736.82	2813.98	0.00	
1745.84	30603.59	0.05	
1762.86	9880.00	0.02	
1768.90	4863.69	0.01	
1775.85	10739.00	0.02	
1792.84	6308.45	0.01	
1804.87	7722.83	0.01	
1810.89	1628.17	0.00	
1827.28	4661.92	0.01	
1837.89	62294.74	0.10	
1848.96	10499.00	0.02	
1859.88	11977.00	0.02	
1864.92	16571.89	0.03	
1884.56	2165.04	0.00	
1908.00	16771.00	0.03	
1924.03	241242.18	0.39	
1946.02	26325.78	0.04	
1961.98	6915.44	0.01	
1972.91	9992.11	0.02	

1978.97	16755.34	0.03
1990.00	146487.19	0.24
1999.99	15255.00	0.02
2012.01	26519.00	0.04
2023.00	19681.00	0.03
2044.51	5771.25	0.01
2056.98	4082.77	0.01
2078.03	2600.41	0.00
2081.14	4421.00	0.01
2086.32	3015.78	0.00
2096.09	197143.92	0.32
2118.06	15606.32	0.03
2134.03	5033.03	0.01
2140.79	5693.38	0.01
2179.10	16817.00	0.03
2197.10	12063.00	0.02
2234.12	2926.96	0.00
2250.10	91004.84	0.15
2253.15	92029.42	0.15
2268.18	586200.16	0.95
2272.19	127853.33	0.21
2282.16	33936.50	0.05
2285.17	26757.86	0.04
2290.14	194800.10	0.31
2295.12	9070.99	0.01
2300.17	66366.28	0.11
2306.11	87697.06	0.14
2312.12	30167.27	0.05
2314.16	8736.82	0.01
2320.14	4138.82	0.01
2322.12	15656.00	0.03
2328.09	12549.57	0.02
2332.25	6948.43	0.01
2345.09	17603.00	0.03
2354.13	4253.72	0.01
2376.19	10236.00	0.02
2404.15	10128.82	0.02
2422.15	11588.24	0.02
2426.18	21289.20	0.03
2429.25	4538.47	0.01
2443.20	135767.60	0.22
2446.20	38194.28	0.06
2465.13	2326.86	0.00
2479.20	2451.49	0.00
2505.27	4400.28	0.01
2509.37	7638.94	0.01
2516.36	2682.93	0.00
2531.24	30814.83	0.05
2534.25	27650.76	0.04
2549.29	421126.78	0.68

2554.28	29298.86	0.05
2561.21	5219.44	0.01
2565.26	170473.60	0.28
2571.22	78890.59	0.13
2575.19	13249.66	0.02
2581.27	9223.94	0.01
2587.20	41033.42	0.07
2593.23	6963.07	0.01
2597.25	12876.55	0.02
2604.28	9760.94	0.02
2629.29	8827.00	0.01
2639.49	23727.13	0.04
2655.37	8734.00	0.01
2673.33	8989.00	0.01
2685.29	6964.80	0.01
2701.31	7797.00	0.01
2702.31	8949.00	0.01
2727.42	7097.00	0.01
2734.37	2395.28	0.00
2738.32	6430.68	0.01
2753.32	1714.70	0.00
2796.46	1770.92	0.00
2807.38	4594.00	0.01
2821.41	548.09	0.00
2838.39	995.32	0.00
2842.49	2057.03	0.00
2850.42	3696.00	0.01
2866.31	595.62	0.00
2900.36	805.09	0.00
2919.49	1661.18	0.00
2980.53	2107.56	0.00
2996.48	1356.24	0.00
3002.58	1299.29	0.00
3008.53	3311.30	0.01
3023.47	10154.01	0.02
3040.46	124890.61	0.20
3062.39	5188.84	0.01
3078.40	1977.70	0.00
3094.52	1040.30	0.00
3110.65	3451.28	0.01
3152.63	4700.00	0.01
3176.56	1706.77	0.00
3191.65	10092.68	0.02
3194.71	2623.96	0.00
3202.62	8355.00	0.01
3208.66	94636.81	0.15
3212.57	7683.62	0.01
3226.32	1414.54	0.00
3232.57	932.68	0.00
3242.62	5450.00	0.01

3250.54	3883.00	0.01
3265.56	3443.00	0.01
3344.75	3056.00	0.00
3360.72	2765.00	0.00
3377.31	2120.74	0.00
3387.86	3439.00	0.01
3404.67	2668.00	0.00
3416.74	2273.00	0.00
3426.72	4294.00	0.01
3457.78	2103.00	0.00
3473.82	2363.65	0.00
3496.71	2015.00	0.00
3508.73	1769.00	0.00
3522.78	1857.00	0.00
3536.78	2033.00	0.00
3551.62	1688.00	0.00
3552.67	1737.00	0.00
3572.71	1735.00	0.00
3582.89	1494.00	0.00
3600.80	1628.00	0.00
3607.94	1948.00	0.00
3623.97	550.01	0.00
3634.95	1798.00	0.00
3642.13	531.77	0.00
3652.85	1826.00	0.00
3666.96	1801.00	0.00
3680.96	1678.00	0.00
3698.92	2264.99	0.00
3715.98	2052.00	0.00
3735.91	2064.00	0.00
3746.81	1694.00	0.00
3755.03	1630.00	0.00
3780.56	2112.00	0.00
3791.02	1520.97	0.00
3796.04	571.81	0.00
3822.96	533.16	0.00
3845.00	1739.00	0.00
3860.22	553.22	0.00
3877.77	1928.64	0.00
3923.19	482.56	0.00
3943.28	758.71	0.00
3966.14	794.63	0.00
3968.17	3009.88	0.00
3979.13	994.25	0.00
3994.30	1948.00	0.00
4021.26	2322.00	0.00
4043.20	2100.00	0.00
4056.18	2068.00	0.00
4111.22	776.02	0.00
4126.32	563.45	0.00

4144.16	478.41	0.00
4162.27	29110.00	0.05
4184.21	2644.00	0.00
4206.03	2580.00	0.00
4273.34	1310.00	0.00
4289.56	1108.00	0.00
4297.44	1405.00	0.00
4313.37	1397.00	0.00
4340.54	1614.00	0.00
4350.45	1107.00	0.00
4365.62	1364.00	0.00
4408.46	1648.00	0.00
4414.49	2017.00	0.00
4424.58	633.38	0.00
4449.48	1081.00	0.00
4465.51	1279.00	0.00
4493.53	965.00	0.00
4535.72	553.55	0.00
4595.57	1794.00	0.00
4610.81	776.26	0.00
4627.66	1131.00	0.00
4644.37	1091.00	0.00
4692.81	541.36	0.00
4708.88	5853.44	0.01
4790.97	1727.00	0.00
4817.00	635.74	0.00
4833.75	1486.00	0.00
4847.95	1393.00	0.00
4856.38	377.93	0.00
4873.77	1161.00	0.00
4887.94	1358.00	0.00
4901.18	866.00	0.00
4918.35	1043.00	0.00
4931.83	1024.00	0.00
4971.23	969.00	0.00
4992.94	964.00	0.00
5029.92	909.00	0.00
5033.34	1540.00	0.00
5065.05	1229.00	0.00
5098.32	168.19	0.00
5113.53	999.00	0.00
5137.23	1292.00	0.00
5151.23	1274.00	0.00
5176.26	1388.00	0.00
5231.16	1030.00	0.00
5282.30	1001.00	0.00
5291.59	820.00	0.00
5309.09	161.76	0.00
5361.43	815.00	0.00
5377.91	941.00	0.00

5443.11	598.00	0.00
5477.38	851.00	0.00
5507.06	891.00	0.00
5590.52	990.00	0.00
5647.61	2365.00	0.00
5661.93	1907.72	0.00
5759.33	969.00	0.00
5791.66	832.00	0.00
5798.52	797.00	0.00
5813.75	939.00	0.00
5830.14	699.11	0.00
5892.03	969.00	0.00
5966.57	688.00	0.00
6008.12	806.00	0.00
6007.01	685.00	0.00
6158.56	604.00	0.00
6688.05	563.00	0.00
6698.72	578.00	0.00
6775.68	551.00	0.00
6778.28	522.00	0.00
6951.44	537.00	0.00
7089.48	541.00	0.00
7136.19	561.00	0.00
7151.38	816.00	0.00
7339.90	561.00	0.00
7342.40	632.00	0.00
7544.56	660.00	0.00
7559.28	94.45	0.00
7659.44	564.00	0.00
7804.92	733.00	0.00
7809.19	962.00	0.00
8091.21	576.00	0.00
8094.03	614.00	0.00