



**University of
Nottingham**
UK | CHINA | MALAYSIA

**Structural studies of Von Willebrand
Factor regulators for Thrombotic
Thrombocytopenic Purpura.**

Zoë Markham-Lee

Thesis submitted to the University of Nottingham
for the degree of
Doctor of Philosophy

School of Pharmacy
March 2024

Supervisors
Prof. Jonas Emsley
Dr Ingrid Dreveny
Prof. Neil Morgan

Abstract

Blood clotting requires an urgent and efficient response with von Willebrand factor (VWF) playing a major role through recruitment of platelets forming a haemostatic plug. VWF is a large multi-domain glycoprotein, with the A1 domain for platelet binding via Gp1b α and the A2 domain containing the cleavage site to reduce multimer size. VWF is regulated by A Disintegrin-like And Metalloprotease with Thrombospondin type I repeats, member 13 (ADAMTS13) to maintain the delicate haemostatic thrombotic balance. ADAMTS13 is a highly specific protease, only cleaving VWF at the scissile bond under specific flow conditions. The process by which this regulation is maintained is still unclear, with details on ADAMTS13 latency and subsequent interaction and cleavage of VWF still to be elucidated.

Thrombotic thrombocytopenic purpura TTP is the clinical deficiency of ADAMTS13 and is a life-threatening thrombotic disorder. The disease is characterised by excessive clotting in the microvasculature and has ~90% mortality if untreated. Caplacizumab is a nanobody recently approved for treatment of TTP, and the structure was resolved in complex with the VWF A1 domain. The VWF A1 domain consists of the main domain region and two flanking autoinhibitory modules termed the NAIM and CAIM, which inhibit A1 activation through disruption of Gp1b α binding. Although mortality rates have subsequently improved, the use of caplacizumab is associated with an increased risk of bleeding so alternative treatment options must be explored to improve patient outcomes. This thesis aims to investigate the structures of ADAMTS13 and VWF A2 domain, as well as VWF A1 domain and the nanobody ND6 to better understand VWF regulation and subsequently inform treatment of thrombotic diseases.

Analysis of TTP patient mutations identified a lack of associations between genotype and phenotype information highlighting the need for a novel ADAMTS13-VWF structure to interrogate the activation cycle of the proteins. Computational studies utilising molecular modelling and docking enabled design of novel VWF constructs, with expression and purification of SUMO-tagged VWF A2 domain fragments as well as a fusion of 2GKG tag to VWF A2 fragments. Initial characterisation of these constructs revealed that, in preliminary binding analysis, the VWF Y1605C mutant

showed a stronger binding affinity to ADAMTS13 than wildtype VWF (KD of 42nm and 450nm respectively), as well as successful complex formation visible utilising size exclusion analysis. This characterisation will help progress understanding of VWF regulation as well as ADAMTS13 structure in normal and disease states following resolution of the crystal structure. Furthermore, the successfully resolved structure of VWF A1 domain in complex with a novel nanobody ND6, gives detailed interface information and a unique conformation of VWF not elucidated in previous structures. The O-glycosylation of T1468 and S1263 is visualised in this novel structure, as well as the interface between both the NAIM and CAIM of VWF A1 occurring from the main body of the domain. Together this information will help deepen the understanding on the fundamental biology of TTP and explores ways to harness this knowledge for improvement in thrombosis disease treatments.

Acknowledgements

I would like to thank Prof. Jonas Emsley for the opportunity to undertake this PhD and providing endless wisdom and support. I have learnt so much through the exposure to multiple projects and collaborations, so I am immensely grateful for all the opportunities including mentoring of students and the publication of multiple papers. I also would like to thank Prof. Neil Morgan for his continued support since my first week as an undergraduate student, helping me to grow as a researcher as well as suggesting this project for me.

A special thanks to Dr Ingrid Dreveny for providing support and feedback in lab meetings and generally keeping the lab running smoothly and making it a lovely place to complete this PhD. Thank you to Dr Monika Pathak for the training and support I very much needed, when starting fresh from my undergraduate degree in the middle of the pandemic. A special thank you also goes to Dr Sigrun Maurer, Dr Matthias Mayer and Dr Jack Yule for the training and support with crystallisation, protein purification and cell culture- your guidance was invaluable to navigating the initial stages of this project. A big thank you to all members of the crystallography lab for all the laughs, sweet treats and support for riding the highs and lows of completing this PhD. I couldn't have picked better colleagues to work alongside.

I'd like to acknowledge all the staff at the BDI for keeping the building running in testing times, as well as the staff at Diamond Light Source for enabling me to produce data worthy of publication and the technical support for all manner of crystallography related questions. A thank you also to my collaborators Prof. Renhao Li and Prof. James Crawley for allowing me to work on such exciting projects.

On a personal note, thank you to my mum and dad for always believing in me and always asking me what I have been up to, even when they didn't understand what I was doing and still being proud regardless; and my sisters for always making me smile when I needed a distraction. Thank you to Megan Trusler for listening to and living by proxy through the rollercoaster that this PhD project has been, and always offering advice or just a shoulder to cry on when I needed it the most. Finally, I would like to thank all my family for their love and support over the past 3.5 years. To Grandad Jeff who would have loved to be able to read this thesis and ask me endless interesting questions, I miss you and this thesis is dedicated to you.

Table of Contents

Abstract	I
Acknowledgements	III
List of papers and meetings	VII
List of abbreviations.....	VIII
List of Figures	IX
List of Tables.....	X
1. Introduction	1
<i>1.1 Haemostasis</i>	<i>1</i>
<i>1.2 ADAMTS13 overview.....</i>	<i>4</i>
1.2.1 Overview.....	4
1.2.2 Synthesis and secretion.....	4
1.2.3 ADAMTS family	5
<i>1.3 Thrombotic Thrombocytopenic Purpura</i>	<i>5</i>
1.3.1 Introduction.....	5
1.3.2 Classification.....	6
1.3.3 Clinical and laboratory features	7
1.3.4 Prediction of disease progression/severity.....	8
1.3.5 TTP treatments.....	8
<i>1.4 Von Willebrand Factor</i>	<i>11</i>
1.4.1 Overview.....	11
1.4.2 VWF synthesis and secretion.....	11
1.4.3 VWF activation and unfolding.....	11
1.4.4 A1 domain.....	12
1.4.5 A2 domain.....	13
1.4.6 VWF in disease.....	14
<i>1.5 ADAMTS13 domain structure.....</i>	<i>15</i>
1.5.1 Metalloprotease domain.....	16
1.5.2 Disintegrin domain.....	17
1.5.3 Cysteine-Rich domain.....	18
1.5.4 Spacer Domain.....	19
1.5.5 Thrombospondin type-1 repeat domains.	21
1.5.6 CUB Domains.....	22
<i>1.6 ADAMTS13 activation.</i>	<i>22</i>
1.6.1 ADAMTS13 allosteric activation by VWF interactions.....	22
1.6.2 Latency of ADAMTS13	23
<i>1.7 Structure prediction and crystallisation</i>	<i>26</i>
1.7.1 AlphaFold for structure prediction.....	26
1.7.2 Previous crystallisation attempts of VWF and ADAMTS13.....	27
<i>1.8 Aims of thesis:.....</i>	<i>27</i>
2. Methods:	29
2.1 <i>Analysis of ADAMTS13 variant data.....</i>	29
2.1.1 Analysis of ADAMTS13 variants in the Ensembl database	29

2.1.2 Analysis of ADAMTS13 mutations in the Ensembl database.....	29
2.2 <i>Designing constructs for VWF-ADAMTS13 interaction studies.</i>	29
2.2.1 Using AlphaFold for prediction of ADAMTS13 and VWF structures.....	29
2.2.2 Guided docking of ADAMTS13 and VWF	30
2.2.3 Sequencing and amplifying DNA constructs.....	30
2.2.4 Site-directed mutagenesis of 6xHis-SUMO-VWF-73	31
2.2.5 VWF-cloning 2GKG.....	33
2.3 <i>Protein expression</i>	34
2.3.1 <i>E.Coli</i> transformation and glycerol stock generation	34
2.3.2 <i>E.Coli</i> protein overexpression.....	34
2.3.3 Drosophila S2 protein expression system (transfection and overexpression).....	34
2.3.4 Calcium phosphate transfection into S2 cells	35
2.3.5 Drosophila Protein overexpression	36
2.4 <i>Protein purification</i>	36
2.4.1 ADAMTS13 MDTCS E225Q construct.....	36
2.4.2 VWF-73 constructs	37
2.4.3 ADAMTS13 DTCS-CUB.....	37
2.5 <i>Complementary techniques for construct characterisation</i>	37
2.5.1 Mass Spectrometry.....	37
2.5.2 Western blot.....	38
2.5.3 Expression and purification of SUMO-protease ULP-1	38
2.5.4 SUMO cleavage of 6xHis-SUMO-VWF73 constructs.....	39
2.5.5 WT ADAMTS13 cleavage assay on VWF.....	39
2.5.6 ITC and SPR	39
2.5.7 Complexing for size exclusion analysis.....	40
2.5.8 Crystallisation trials	40
2.6 <i>Techniques for crystallisation and characterisation of VWF-ND6 complex</i>	40
2.6.1 Crystal growth and X-ray crystallography data collection	40
2.6.2 Generation of model for structure solution.	41
2.6.3 Data processing and model building with X-ray crystallography data.....	41
2.6.4 Aligning for comparison of complex structures.	43
2.6.5 Analysis of interactions.....	43
3. Linking genotype and phenotype: exploring the TTP mutations, and the structural effect on ADAMTS13.....	44
3.1 <i>Introduction</i>	44
3.2 <i>Results</i>	46
3.2.1 Summary of themes in ADAMTS13 mutations in the Ensembl database.....	46
3.2.2 Assessing the link between genotype and phenotype in TTP patients	50
3.3 <i>Discussion</i>	59
4. Prediction of VWF-ADAMTS13 interactions to inform construct design for structural studies.	62
4.1 <i>Introduction</i>	62
4.2 <i>Results</i>	64
4.2.1 Comparison of experimental and predicted ADAMTS13 structures.....	64
4.2.2 Guided docking of ADAMTS13 and VWF A2 domain	67
4.2.3 Design of VWF mutants to improve stable interaction with ADAMTS13.	69
4.2.4 Generation of 6xHis-SUMO-VWF73 mutants for complexing with ADAMTS13.....	71
4.2.5 2GKG construct design.....	72
4.2.6 Generation of 2GKG-VWF constructs for complexing with ADAMTS13.....	74
4.2.7 ADAMTS13 construct design for crystallisation studies.	74
4.3 <i>Discussion</i>	76

5. Exploring complex formation and the interactions between ADAMTS13 MDTCS and VWF A2 domains.....	80
5.1 Introduction	80
5.2 Results.....	82
5.2.1 Protein expression and purification	82
5.2.1.1 VWF constructs	82
5.2.1.2 MDTCS E225Q mutant	89
5.2.2 Conformation of successful purification with western blot analysis and Mass spectrometry.....	90
5.2.3 Expression and purification of SUMO protease ULP1 for SUMO cleavage of 6xHis-SUMO-VWF73 constructs.....	91
5.2.4 ADAMTS13 and VWF cleavage assays.....	92
5.2.5 Isothermal Calorimetry and Surface Plasmon Resonance for analysis of binding energetics between ADAMTS13 and VWF.....	93
5.2.6 Complexing MDTCS with VWF	95
5.2.7 Crystallisation trials	98
5.3 Discussion.....	98
6. Characterising VWF and novel nanobody ND6.....	103
6.1 Introduction	103
6.2 Results.....	105
6.2.1 X-ray diffraction experiments	105
6.2.2 Crystal packing and space group determination	107
6.2.3 Differences between subunits in the asymmetric unit (ASU).....	107
6.2.4 A1-ND6 interface interactions	109
6.2.5 Structural differences of ND6 bound AIM-A1, to unbound AIM-A1 and caplacizumab bound AIM-A1 structures.....	114
6.2.6 O-glycosylation of AIM-A1	116
6.2.7 ND4 high resolution structure.....	116
6.2.8 AIM-A1 crystallisation trials	117
6.3 Discussion.....	118
7. Discussion and conclusions	122
7.1 Study limitations and future work.....	126
7.2 Concluding remarks.....	127
8. Supplementary data	129
References	134

List of papers and meetings

Publications:

Markham-Lee Z, Morgan NV, Emsley J. Inherited ADAMTS13 mutations associated with Thrombotic Thrombocytopenic Purpura: a short review and update. *Platelets*. 2023;34(1):2138306.

<https://doi.org/10.1080/09537104.2022.2138306>

Soukarieh, F., Mashabi, A., Richardson, W., Oton, E. V., Romero, M., Dubern, J. F., Robertson, S. N., Lucanto, S., **Markham-Lee, Z.**, Sou, T., Kukavica-Ibrulj, I., Levesque, R. C., Bergstrom, C. A. S., Halliday, N., Kellam, B., Emsley, J., Heeb, S., Williams, P., Stocks, M. J., & Cámara, M. (2024). Design, Synthesis, and Evaluation of New 1*H*-Benzo[*d*]imidazole Based PqsR Inhibitors as Adjuvant Therapy for *Pseudomonas aeruginosa* Infections. *Journal of medicinal chemistry*, 67(2), 1008–1023. <https://doi.org/10.1021/acs.jmedchem.3c00973>

Arce, N. A., **Markham-Lee, Z.**, Liang, Q., Najmudin, S., Legan, E. R., Dean, G., Su, A. J., Wilson, M. S., Sidonio, R. F., Lollar, P., Emsley, J., & Li, R. (2024). Conformational activation and inhibition of von Willebrand factor by targeting its autoinhibitory module. *Blood*, 143(19), 1992–2004.

<https://doi.org/10.1182/blood.2023022038>

Meetings:

The interplay between ADAMTS13 and VWF and their role in disease. Poster Presentation at University of Nottingham, School of Pharmacy research day 2023.

The complex structure of Von Willebrand Factor AIM-A1 domain and ND6 nanobody reveals novel A1 domain conformation through autoinhibitory module stabilisation. Poster presented at CCP4 Northern structural biology meeting and MRC AIM/IMPACT symposium 2024.

Conformational allosteric inhibition of von Willebrand factor by targeting its A1 domain autoinhibitory module with a novel nanobody. Dr Shabir Najmudin (University of Nottingham, UK)- Results of this thesis presented on my behalf by Dr Najmudin at BSHT Annual Scientific Meeting 2024 (Oral communication).

List of abbreviations

ADAMTS13	a disintegrin-like and metalloprotease with thrombospondin type 1 repeat motifs member 13
AIM	Autoinhibitory modules
ASU	asymmetric unit
CCP4	Collaborative Computational Project No. 4
Cys	Cysteine-rich domain
CDR	Complementarity-determining region
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHAPSO	3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate
CHD	Cholic acid, 3 α ,7 α ,12 α -trihydroxy-5 β -cholan-24-oic acid
CUB domain	C1r/C1s, Uegf sea urchin fibropellins, and bone morphogenic protein 1 domain
Dis	Disintegrin-like domain
DTCS	ADAMTS13 disintegrin-like, thrombospondin repeat-1, Cystine-rich, spacer, TSP2-8 and CUB domains
FFP	Fresh frozen plasma
GOF	Gain-of-function
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethane-sulfonic acid
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
LBD	Ligand-binding domain
MBP	Maltose-binding protein
MDTCS	ADAMTS13 metalloprotease, disintegrin-like, thrombospondin repeat-1, Cystine-rich, spacer domains
MP	Metalloprotease domain
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SPR	Surface Plasmon Resonance
TSP	thrombospondin
TTP	Thrombotic thrombocytopenic purpura
VWF	Von Willebrand factor

List of Figures

- 1.1 VWF mature protein domain architecture
- 1.2 Schematic of ADAMTS13 mature protein domain structure
- 1.3 Molecular surface representation of ADAMTS13 MDTCS and VWF A2 domains.
- 1.4 Molecular visualisation of ADAMTS13 metalloprotease domain features.
- 1.5 Schematic of the domain architecture of ADAMTS13.
- 1.6 Proposed model of binding between the ADAMTS13 Spacer and CUB domains
- 2.1 Schematic representation of VWF A2 constructs.
- 3.1 Graphical representation of ADAMTS13 mutations (location, mutation type and variant type)
- 3.2 Graphical depiction of ADAMTS13 mutations (SIFT and REVEL)
- 3.3 ADAMTS13 mutations by domain location
- 4.1 AlphaFold prediction for ADAMTS13.
- 4.2 Aligned structures of ADAMTS13 experimental and predicted structures.
- 4.3 Guided docking between MDTCS E225Q and VWF A2 domain fragment.
- 4.4 Modelling and docking of VWF mutant constructs.
- 4.5 Cartoon representation of the VWF A2 domain comparing WT and G1629E mutant constructs.
- 4.6 AlphaFold modelled structures of VWF-2GKG constructs docked with ADAMTS13 domains.
- 4.7 AlphaFold predictions of DTCS and comparison to MDTCS experimental structure and predicted full-length ADAMTS13 structure.
- 5.1: Purification of 6xHis-SUMO-VWF constructs
- 5.2: Purification of 2GKG-VWF constructs
- 5.3: Purification of MDTCS E225Q construct
- 5.4: Western blotting to confirm purification of poly-histidine tagged constructs.
- 5.5 Purification of ULP-1 protein, and cleavage of SUMO-tagged VWF.
- 5.6 VWF constructs cleavage by ADAMTS13.
- 5.7 Binding analysis of MDTCS and VWF constructs by ITC.
- 5.8 Binding analysis of MDTCS and VWF constructs by SPR
- 5.9 Size exclusion analysis to assess complex formation of MDTCS E225Q with 6xHis-SUMO-VWF73
- 5.10 Crystallisation trials of MDTCS and 6xHis-SUMO VWF constructs.
- 6.1: Cartoon of AIM-A1 structure with Caplacizumab and ND6 localised to binding sites
- 6.2: Images of crystals and their diffraction of AIM-A1/ND6 complex
- 6.3 ASU of AIM-A1/ND6 crystal
- 6.4 CPS molecules facilitate the AIM-A1/ND6 crystal formation.
- 6.5 ND6 nanobody cartoon and schematic representation.
- 6.6 VWF AIM-A1 and ND6 complex interface.
- 6.7 VWF AIM-A1 and ND6 cartoon and molecular models of experimental structure.
- 6.8 Cartoon and molecular structures of AIM-A1 domain experimental structures.
- 6.9 AIM structure with electron density map
- 6.10 High resolution ND4 structure with comparison to ND6.
- 6.11: Images of crystals of VWF AIM-A1.

List of Tables

1.1 ADAMTS13-VWF exosites.

2.1 Mutagenesis primers used for generation of mutant 6xHis-SUMO-VWF-73 constructs.

2.2 PCR protocol for mutagenesis reactions.

3.1 ADAMTS13 mutations in TTP patients.

6.1: Crystallographic data collection and refinement statistics.

1. Introduction

1.1 Haemostasis

Haemostasis is the process by which a clot is formed at the site of vessel damage to prevent blood leaking from the vasculature. Haemostasis involves a complex cascade of enzyme-mediated events in multiple steps or pathways. When a blood vessel is damaged, an urgent and efficient response is required to prevent major harm from occurring through excessive bleeding. Primary haemostasis involves the initial formation of a clot to plug the hole in the vessel whereas secondary haemostasis stabilises the clot through deposition of insoluble fibrin (1). Secondary haemostasis can be further broken down into the intrinsic and extrinsic pathways, as well as the common pathway. The intrinsic pathway is activated through exposed endothelial collagen within the blood stream, whereas and the extrinsic pathway is activated external damage and release of tissue factor (2).

VWF is a large glycoprotein responsible for regulating platelet adhesion at sites of vessel damage, and their subsequent aggregation to form blood clots (3). VWF consists of a series of domains, D1-D2-D'-D3-A1-A2-A3-D4-C1-6-CK, with D1 and D2 forming the propeptide; these domain regions are responsible for binding different proteins (Figure 1.1A) (3). Under minimal stress in haemostasis conditions, VWF circulates in a closed globular form. A single monomer can be 70nm in size and the globular structure under stasis is a disordered coil formed of many monomers interacting in a disordered manner (4, 5). Increasing force leads to intermediate partial elongation of VWF, with the fully elongated version of VWF containing many multimers forming 100-1000 μm -long strings (Figure 1.1C) (5).

Following vessel damage and subsequent increased shear flow in the blood, elongated VWF binds to exposed subendothelial collagen through its A3 domain (4). This elongation also reveals the GPIIb α platelet binding site on the A1 domain enabling binding of platelets, facilitating platelet recruitment to the site of vascular injury (5). The D'D3 domains have been shown to inhibit A1 binding to platelet glycoproteins through their close proximity, which is subsequently disrupted under shear stress (6). These initial interactions allow platelets to form low-affinity

interactions with the endothelial cells, before integrins ($\alpha\text{IIb}\beta\text{3}$) then bind to VWF via the C-domains forming a stronger platelet adhesion (7).

To increase the efficiency of this process, larger VWF multimers are often found to be circulating as they have a higher affinity to binding platelets (8). VWF is therefore released and circulates as multimers that can be as large as 100 or 200 mers (but normally 20-40 mers), these so called ultra-large (UL) multimers can be potentially pathogenic and lead to microthrombi formation in small vessels if uncontrolled (9). Regulating the size of these VWF multimers is a process tightly controlled to maintain haemostasis ; a role fulfilled by a disintegrin-like and metalloprotease with thrombospondin type 1 repeat motifs member 13 (ADAMTS13) (3). The ADAMTS13 gene encodes a plasma metalloprotease and is responsible for cleavage of large VWF multimers at the scissile bond (Tyr1605-Met1606). This cleavage prevents excessive ultra-large VWF multimer formation and therefore perturbs increased aggregation of platelets. Thus, maintains the delicate haemostatic thrombotic balance (3).

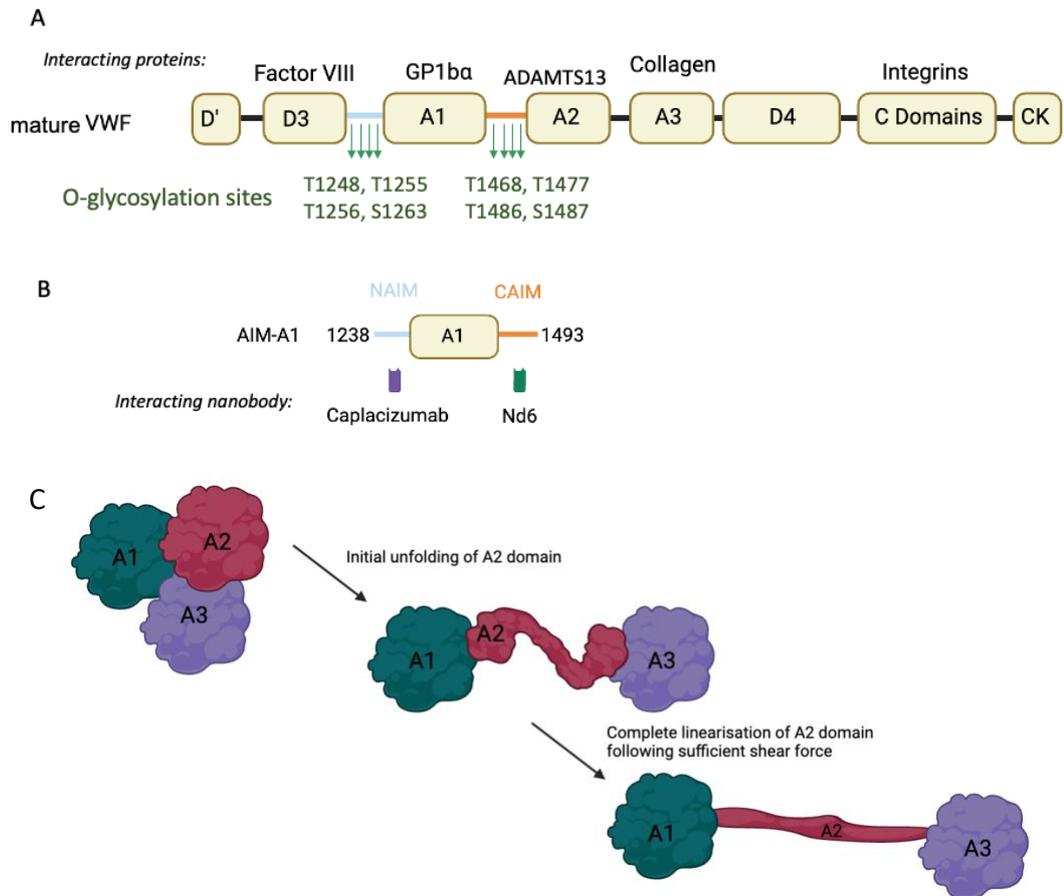


Figure 1.1 VWF mature protein domain architecture (A) Schematic of VWF mature protein domain structure. O-glycosylation sites and respective interacting ligand locations highlighted (B) Schematic structure of the VWF AIM-A1 domain, with flanking autoinhibitory modules and targeting nanobodies highlighted. The NAIM coloured (in blue) consists of residues 1238–1271 the CAIM (in orange) consists of residues 1459–1478. (C) Schematic depicts the unfolding of the VWF A1-A2-A3 domains when VWF undergoes shear stress. Initial interactions with collagen and the A3 domain led to partial unfolding. When the shear stress threshold is reached the Cys1669-Cys1670 plug is removed from the A2 domain allowing full linearisation.

1.2 ADAMTS13 overview

1.2.1 Overview

The *ADAMTS13* gene consists of 29 exons at chromosome location 9q34.2 (10). The ADAMTS13 protein is a metalloprotease, for cleavage of VWF at the scissile bond: cleavage is restricted to this specific site. ADAMTS13, like other Zn²⁺ dependent metalloproteases, contains (from the mature N-terminus) a metalloprotease (MP), disintegrin-like (Dis), thrombospondin type 1 (TSP) repeats, cysteine-rich (Cys) and spacer domains - together these domains are termed MDTCS (Figure 1.2) (11). ADAMTS13 also contains 8 TSP repeats as well as two C1r/C1s, Uegf sea urchin fibropellins, and bone morphogenic protein 1 (CUB) domains at the C-terminus which are not found in any other ADAMTS protein (Figure 1.2) (12, 13).

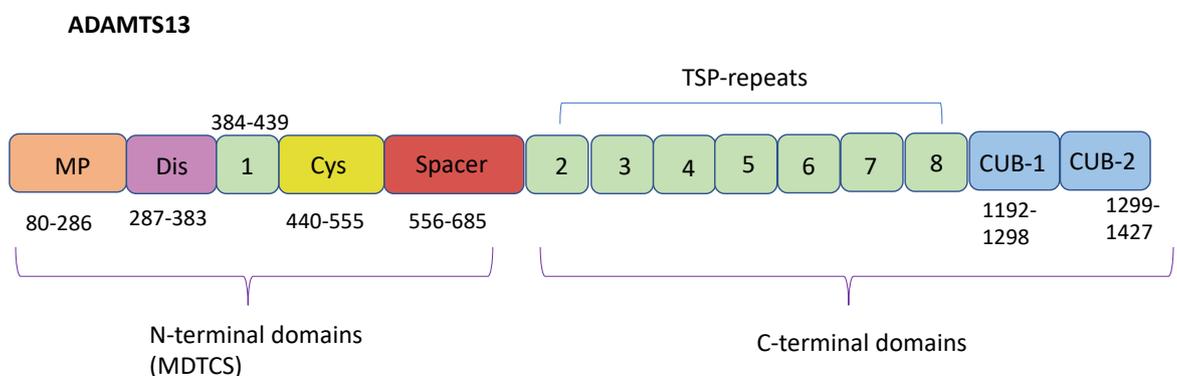


Figure 1.2 Schematic of ADAMTS13 mature protein domain structure, with residue numbers included. MDTCS refers to the N-terminal region of the protein domains and the TSP-repeats (numbered within the schematic domain shapes) and CUB domains make up the C-terminal region of ADAMTS13.

1.2.2 Synthesis and secretion

ADAMTS13 synthesis occurs primarily in hepatic stellate cells of the liver, with evidence of a lesser production in platelets, megakaryocytes and vascular endothelial cells (14-17). N-glycosylation and O-fucosylation of ADAMTS13 occurs in the endoplasmic reticulum and both steps are essential for secretion of ADAMTS13 (18). Following this, conversion of N-glycans to complex types occurs in the Golgi apparatus and subsequently aids in VWF cleavage (19). C-Mannosylation was also found on TSP-1, TSP-4 linker and TSP-8 of ADAMTS13, with mutagenesis studies (mutations of the WXXW motif) indicating reduced secretion; but sufficient functional data is lacking with more research needed to understand the role of this

modification (20, 21). In circulation, concentrations of ADAMTS13 are 0.5–1.4 µg/ml, with an estimated half-life of two to four days (22-24).

1.2.3 ADAMTS family

ADAMTS13 is an ~180-kDa multi-domain metalloprotease part of the ADAMTS family(10). The 19 ADAMTS proteins are all zinc-metalloproteases, consisting of multiple domains which are conserved. Further domains on the C-terminus allow categorisation of ADAMTS proteins into sub-types based on their function. For example, ADAMTS 2, 3 and 14 contain specialised domains for their function as procollagen N-propeptidases. ADAMTS 9 and 20 contain GON-1 domains for aggrecanase/ proteoglycanase activity alongside ADAMTS 1, 4, 5 and 8 which consist only of MDTCS domains (with varying numbers of TSP repeats) (13). Many ADAMTS proteins are secreted as zymogens with the prodomain preventing activation, requiring cleavage to produce the active protein. For example, ADAMTS5 requires furin cleavage at furin-consensus sites in their propeptide to allow activity (25). The unique and unusually short propeptide of ADAMTS13 bears no regulatory control over folding or activity, differing to other ADAMTS family proteins(26, 27).

Genetic disorders can be caused by mutations in ADAMTS proteins with an autosomal recessive pattern as seen with ADAMTS13 and TTP. Specifically, ADAMTS2 mutations cause Ehlers-Danlos syndrome (a connective tissue disorder) and ADAMTS3 mutations cause Hennekam lymphangiectasia-lymphedema syndrome 3 (a defect within the lymphatic system) (28, 29). ADAMTS5 and ADAMTS7 are currently being targeted for treatments of osteoarthritis and atherosclerosis respectively, with therapeutics designed to inhibit both proteins (30, 31). This is in contrast to ADAMTS13 where therapies for TTP aim to replace the loss of function (11).

1.3 Thrombotic Thrombocytopenic Purpura

1.3.1 Introduction

The condition now known as thrombotic thrombocytopenic purpura, TTP, was initially called Moschcowitz syndrome, when first reported by Eli Moschcowitz in

1924. He presented a case of a 16-year-old girl reporting fever, aches and petechiae on the arms; unfortunately, she rapidly declined and after entering a coma died a day later. The post-mortem reported, microthrombi in the small vessels including those of the heart and kidneys which are now known as key clinical features of TTP (32). The condition was subsequently reported by numerous other clinicians including Schulman et al. and Upshaw et al., when the congenital form of disease was known as Upshaw-Schulman syndrome (33, 34).

Currently, prevalence of this rare condition is predicted to have an approximate annual incidence of 10 cases in 1 million people but varies depending on location (35, 36). Key peaks in presentation occur in childhood and adulthood (37, 38). TTP is classified as a medical emergency due to fatality if not promptly treated. Quick diagnosis and efficient treatment have dramatically reduced mortality rates from 90%, when the disease was first discovered, down to 10-15% (39, 40).

1.3.2 Classification

TTP is classified into 2 types, immune-mediated (iTTP) or congenital (cTTP). Both types of disease have ADAMTS13 activity level $< 10\%$, with the acquired iTTP as a result of auto-antibody binding to ADAMTS13 and neutralising the protein through preventing cleavage of VWF, or increasing clearance of the protein (41, 42). Congenital TTP involves a mutation in the ADAMTS13 gene which results in severe protein deficiency; either through diminished or abolished protein secretion or activity (43).

Age of presentation can be a good indicator of the type of TTP, with adult on-set disease far more commonly iTTP, around age 40, compared to infancy onset disease which is indicative of cTTP (36). However, the importance of testing for confirmation of classification is clear, as a large percentage of cTTP are adult-onset, particularly associated with first-pregnancy which can be a trigger of the first acute episode of disease (44). Other triggers of iTTP have also been reported including infections, oestrogen-containing birth control, antiplatelet drugs, and immunosuppressive agents (45-50). Generally, triggers can be summarised as events with a burden on normal haemostasis.

1.3.3 Clinical and laboratory features

Clinical features upon presentation of TTP are extremely varied. Previously the pentad of fever, thrombocytopenia, microangiopathic haemolytic anaemia, neurological and renal issues were used for diagnosis, which has now been strongly disproved due to lack of presentation of all symptoms in many patients (51). Instead any symptoms that could be related to microvascular ischemia in organs such as the heart, brain and kidneys should be suggestive of TTP and hence require laboratory diagnostic testing (52).

Assaying of ADAMTS13 activity can vary between both countries, and even institutions. The principles remain the same between assays, involving a solution of VWF, or short VWF peptide containing the A2 domain scissile site, to which patient plasma is added and the cleavage efficiency is then measured using fluorescence or immunoassay techniques. Each assay has different advantages (including speed and specificity) and importantly results between assays are not always comparable, so it's useful for follow up tests to use the same technique (52). The sequence of assays conducted for diagnosis will be explored further in Chapter 3.1.

Laboratory tests are often conducted initially to give an indication of thrombosis and the potential severity of ADAMTS13 deficiency. The PLASMIC score summarises results from these tests into a scoring system with one point given for scores of: Platelet count $<30 \times 10^9/L$, Parameters of haemolysis (reticulocyte count $>2.5\%$, haptoglobin undetectable, indirect bilirubin $>2 \text{ mg/dL}$), associated conditions (no active cancer, no history of solid-organ or hematopoietic stem cell transplant), Mean corpuscular volume $<90 \text{ fL}$ and international normalised ratio <1.5 . Risk is then calculated from the points for each condition met (0-4 low, 5 intermediate and 6-7 high). It's important to note this score has limited utility in long-term prediction of disease but does help with ensuring treatment is quickly received by high-risk patients, and low-risk disease diagnosis is confirmed before unnecessary treatments are given (53).

1.3.4 Prediction of disease progression/severity

Currently TTP is understood to be a heterogenous disease, with patients presenting with a different combination of clinical and laboratory features of TTP, thus diagnosis and subsequent disease progression prediction can be complicated. For cTTP, exploration of mutations in TTP and their influence on disease severity and progression is an area of research that is still progressing.

New developments into using machine learning processes to predict the effect of mutations on proteins throughout the human genome, are helping to understand the link between genotype and phenotype of diseases. Previously, pathogenicity prediction software such as SIFT and Polyphen rely strongly on sequence conservation estimates meaning results can be inconsistent, and the techniques have been classified with high sensitivity but low specificity in predictions (54-56). It is important to consider the function of proteins and specific domains when comparing sequences, as regions of conservation are often but not always functionally important which could influence pathogenicity prediction. Furthermore, protein structures are also more strongly conserved than sequences which forms the basis of homology modelling for protein structures and is also useful for predicting pathogenic structural changes (57). AlphaMissense instead utilises machine learning and is trained using functional data to provide more accurate pathogenicity predictions. So far improvements have been reported as modest, however with greater training possesses potential to aid in future disease prediction of mutations (58). In order to achieve this potential with prediction software, a good model of protein structure must exist with understanding of latency and interaction, to enable understanding of how this is affected by mutations. The association between genotype and phenotype of TTP will be discussed further in Chapter 3.

1.3.5 TTP treatments

TTP previously had extremely low survival rates until the 1990s, when therapeutic plasma exchange (TPE) was established as the standard of care, which improved survival from 10- 20% up to 80% (35). Plasma therapy remains the standard of care currently and is initiated as soon as TTP is suspected or diagnosed. Treatment is then performed daily until a reduction in disease manifestations are seen, including

stabilisation of blood counts as well as organ-specific issues (cerebral or renal dysfunction). Plasma exchange was shown to be superior over plasma infusion in phase 3 clinical trial of patients with TTP in 1991, whilst also highlighting the need for better understanding of disease pathophysiology to aid to decrease mortality rates further (59). Whilst treatments may initially provide disease remission, a third of patients are reported to suffer relapses with limited information in clinical or demographic features as to why or when relapses will occur (60, 61).

Guidelines were introduced for the consistent reporting of terms relating to relapse by Vesely et al 2003 (62). This includes 'response to treatment' defined by a platelet count above $150 \times 10^9/L$, with 'complete response' also involving normalizing LDH and clinical recovery. Furthermore, recurrent disease within 30 days after reaching treatment response defines an exacerbation, and whereas recurrence 30 days or longer after reaching treatment response is a relapse. Sarode et al., (2013) also proposed that refractory disease is defined by no treatment response by day 30 and/or no durable treatment response by day 60 (63).

New treatments are being researched to improve patient quality of life by reducing the invasive nature of treatments as well as to address issues with allergic reaction or relapse. In 2020, the ISTH published guidelines for the treatment and management of TTP (64). For cTTP this included plasma or factor VIII (FVIII) concentrate infusions which were only recommended for patients in remission if they are pregnant, otherwise the watch-and-wait strategy was recommended. This is in contrast to treatment for iTTP where there are multiple treatment avenues to pursue relating to immune modulating drugs such as rituximab and corticosteroids (64). Another treatment option approved in the US, Europe and Australia amongst other countries is the nanobody caplacizumab (65-67). A nanobody is an antibody fragment (from the heavy chain) consisting of a single monomeric variable domain from the Camelidae family. Caplacizumab is a bivalent humanized nanobody containing 2 identical monovalent building blocks targeting the A1 domain of VWF, the monovalent unit was isolated from a llama immunized with the recombinant A1 domain of VWF and binds to the N-terminus (NAIM) and $\alpha 1\beta 2$ loop located at the bottom face of the A1 domain (68). The mechanism of action is through binding to

the A1 domain of VWF, blocking platelets from binding to VWF and preventing aggregation (69).

Caplacizumab is now recommended for use in the ISTH guidelines, if sufficient monitoring of patients can take place, and was found to improve patient survival in multiple iTTP populations. Previously, data for use of caplacizumab in cTTP patients was lacking, and subsequently was not recognised as a treatment in the guidelines, however research since has addressed this and indicated improvements in patient treatment response (70). More details on caplacizumab, and the improvements being made to nanobody treatments for TTP are explored in Chapter 6.

Attempts to replenish ADAMTS13 function *in vivo* was initially explored utilising constructs of N-terminal ADAMTS13 domains with some efficiency to partially restore some of the abolished proteolytic activity in ADAMTS13^{-/-} murine models (71). The latest novel treatment exploring a direct replacement of the dysfunctional ADAMTS13 is the use of recombinant ADAMTS13 (rADAMTS13) protein, with potential utility for both cTTP and iTTP. ADZYNMA is a purified recombinant protein, expressed in Chinese Hamster Ovary (CHO) cells and features a mixture of two amino acid variations of ADAMTS13 (either Q or R at position 97) the structures of which have not been experimentally determined (72, 73). A new phase III clinical trial of the rADAMTS13 has thus far reported no safety concerns or evidence of acute events whilst providing increased ADAMTS13 activity compared to traditional plasma therapies. The results of this study are yet to be published, and long-term effects are still to be monitored and reported accordingly (74, 75). Progress in development of rADAMTS13 therapy as well as nanobody regulation of VWF provides hope for improvements in TTP patient treatment, which can be further improved by detailed understanding of disease pathophysiology and the direct action of the drugs on ADAMTS13 and VWF to restore haemostasis.

1.4 Von Willebrand Factor

1.4.1 Overview

The VWF gene (~178kb) located on chromosome 12 encodes an adhesive and multimeric glycoprotein. The VWF monomer consists of 52-exons encoding 2813 amino acids (76). Pro-VWF undergoes co- and post-translational modifications, including addition of 16 N- and 10 O-linked glycans (OLGs) that comprise 18.7% of the total molecular weight (77). N-linked glycosylation has been extensively mapped on VWF, and subsequent modulation of ADAMTS13 interaction through mutagenesis studies of N-glycosylation sites (78, 79). Alternatively, O-linked glycosylation influences the structure of VWF by altering accessibility of the A1 domain for platelet binding (77).

Both the N- and O- linked glycosylation of VWF is also involved in the blood group determination, through covalently linked ABO(H) blood group determinants as residues on both the *N*- and *O*-glycan termini. These blood groups can also reflect minor differences in VWF levels with O-group reported to have lowest VWF levels with groups A, B and AB with increasing VWF levels respectively (80).

1.4.2 VWF synthesis and secretion

VWF is synthesised in both megakaryocytes and endothelial cells. In endothelial cells biogenesis initially involves dimerisation and signal peptide cleavage in the endoplasmic reticulum (81). Subsequently, VWF forms long disulphide linked concatemers consisting of head to tail multimer formation, for which the mechanism is partially elucidated, requiring the acidic environment of the WPBs (82). At this point the propeptide (required for the intracellular processing) is also cleaved, allowing VWF to enter the secretory pathways of the cells (83). These large VWF tubule structures are essential for haemostasis with availability of a large number of platelet binding sites (8).

1.4.3 VWF activation and unfolding

Whilst under stasis VWF circulates in a coiled structure with irregular packing but once under shear flow above the shear threshold activation is triggered, elongation

can be seen in experiments above 5000 s^{-1} with single platelets adhering (84). Under shear flow during one rotation VWF follows 2 cycles of elongation and contraction with force transmitted through domains onto neighbours as well as O-glycans (85). It is this process that exposes domain regions that were previously inaccessible to allow binding of ligands to trigger the full activation and unfolding of VWF (As seen in A1-3 domains in Figure 1.1C) (5).

1.4.4 A1 domain

The A1 domain of VWF remains in the globular formation under normal blood stasis, however following shear flow experiences shear induced changes. AIM-A1 is the term used to describe the A1 domain and the flanking autoinhibitory modules (AIMs) found at both C and N- termini (termed CAIM and NAIM respectively) (Figure 1.1B). The A1 domain contains the important GPIIb/IIIa binding site crucial for platelets binding and control and generation of haemostatic plug. The interactions between VWF A1 and GPIIb/IIIa have been studied intensively, with a crystal structure available of their interaction (1SQO (86)). GPIIb/IIIa instigates structural changes through its LBD at the alpha1-beta2 loop of A1, as well as the GPIIb/IIIa Beta-switch region to beta2/3 strands of A1 as described in Arce et al 2021 (68, 86).

There is still some debate about the role of the CAIM and NAIM and the effect they have on the inhibition of GPIIb/IIIa binding to A1, a summary of the arguments are presented in Bonazza et al 2022 (87). However recent evidence and investigation suggests the AIMs are co-operatively responsible for prevention of activation of VWF unless at high shear force. This concept is supported by previous structural studies from Arce et al (2021) with nanobodies targeted to the NAIM and CAIM (Caplacizumab and ND6 respectively) used to investigate inhibition of the AIMs, and the research is continued in this thesis (and published in Arce et al 2024) (68, 88). The AIMs adopt a conformation with some stability that requires forces matching 20pN to disrupt the confirmation of the AIMs but not the A1 domain entirely due to the strong disulphide bond present between Cys1272 and Cys1458 requiring higher forces still (89). Exploration of the structure and function of the AIMs will be explored further in Chapter 6.

The A1 domain of VWF also contains the heparin and ristocetin binding sites which are utilised for studies of haemostasis. Ristocetin is a glycopeptide used to activate VWF via the A1 domain, the activation is artificial but is closely related to physiological conditions (90). The VWF-ristocetin assay is commonly used for characterising the effect of mutations of VWF (resulting in von Willebrand disease) by simulating the effects of VWF- GPIIb α interactions (91, 92). The binding site at residues E1463–D1472 (specifically the proline rich sequence) in the CAIM has been identified as crucial for GPIIb α binding mediated by ristocetin (93).

The A1 domain also has an inhibitory effect on the cleavage of VWF in the A2 domain under normal physiological conditions. When GPIIb α binds under shear stress this alleviates the inhibition, creating feedback from platelet adhesion allowing cleavage of VWF in the A2 domain to prevent excessive thrombi size (94).

1.4.5 A2 domain

Cleavage of VWF is critical for reducing VWF multimer size, thus ensuring thrombi are not too large and subsequently occlude blood vessels. The A2 domain contains the important site of the scissile bond, Y1605-M1606, where ADAMTS13 cleavage occurs. The positioning of VWF residues in the MP domain subsites has been predicted utilising both docking and mutagenesis experiments (Figure 1.3). P1 residue Y1605 and P1' residue M1606 both interact with ADAMTS13 metalloprotease (MP) domain residues forming the S1 pocket (V195 and L151) for P1, and S1' pocket (D252-256) for P1' (95). The residues around this VWF scissile bond are also critical for optimal cleavage by ADAMTS13. L1603 (the P3 residue), is important for orientation of the scissile bond at the ADAMTS13 active site and binds at the P3 pocket (L198, 232 and 274) (96, 97). This evidence for VWF binding to ADAMTS13 relies exclusively on a latent protein information including the crystal structure of ADAMTS13 (PDB: 6QIG) so the location of the A2 domain and importance of peripheral VWF residues affecting activation of the ADAMTS13 MP domain are yet to be fully understood.

The A2 domain shows an evolutionary structural difference to other VWF A domains with loss of the α 4-helix, that both allows unfolding and slower refolding, among other features to allow the A2 domain to act as a shear sensor (98). Within

the A2 domain, C-terminal cysteine residues (C1669 and 1670) form a disulphide bond that interacts with the hydrophobic core of the domain; this interaction provides the stability for the domain and must first be removed to allow for further unfolding. It's thought this region initially prevents aberrant cleavage of VWF, whilst also allowing complete unfolding of the A2 domain to align with the ADAMTS13 active sites (99). Furthermore, the α 3- β 4 loop with a Ca^{2+} binding site was previously investigated to contribute to the unfolding of the A2 domain (98, 100). This is in comparison to the A1 and A3 domain either side that remain largely globular (9) (Figure 1.1C).

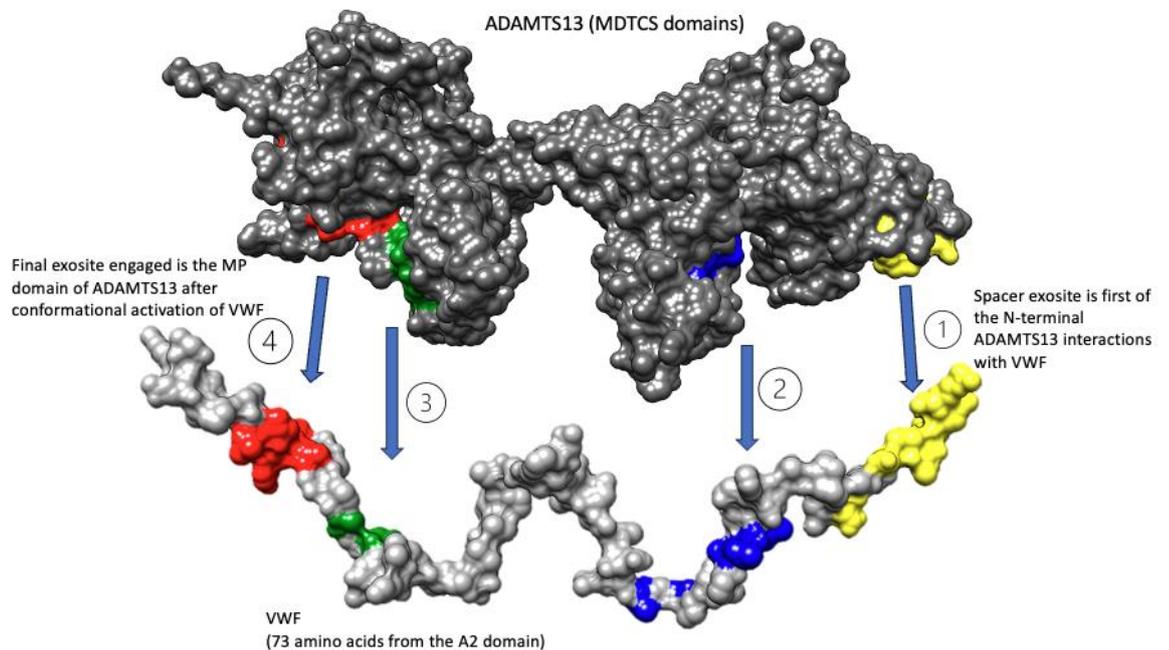


Figure 1.3 Molecular surface representation of ADAMTS13 MDTCS domains (from PDB: 6QIG) and linear VWF A2 (73 amino acids) with active site and exosite interactions coloured (Red for metalloprotease domain, green for disintegrin-like, blue for cysteine-rich, yellow for spacer). The order of discontinuous exosite interactions between ADAMTS13 and VWF are indicated.

1.4.6 VWF in disease

Mutations of any coagulation factor, and subsequent reduction in activity, can lead to a plethora of genetic bleeding disorders involving aberrant blood clot formation. VWF disease (VWD) encompasses multiple disorders of differing severity and issues with the VWF protein, with tests to differentiate between types (101). Plasma

levels of VWF can range between 0.6–2 U/mL with lower levels indicative of disease, at what point reduction indicates disease is currently still debated (102, 103). A “low VWF” diagnosis is often given to patients with levels between 0.3 IU/ml and 0.5IU/ml (104). Type 1 VWD is a mild to moderate disease with partial deficiency of normal VWF protein, and is the most common whereas Type 3 is a severe disease with near complete deficiency of VWF protein (105).

Type 2 VWD consists of abnormal VWF protein resulting in aberrant haemostasis. Type 2 can be further broken down into subtypes 2A (decreased platelet interaction, affecting A2 domain) 2B (increased affinity to GPIb α , affecting the A1 domain), 2M (decreased affinity to platelets, affecting A1 and A3 domains) and 2N (decreased affinity for Factor VIII, affecting the D'-D3 region). Type 2A refers to variants of VWF, exhibiting decreased platelet adhesion and presence of smaller VWF multimers (106). Mutations in the A2 domain can affect protein secretion or alternatively leave the protein more susceptible to cleavage by ADAMTS13 (even without injury). The resultant conditions involve excessive bleeding symptoms including epistaxis, delayed clotting response and excessive bruising (107). Type 2B mutations found in the A1 domain (responsible for binding to GPIb α) enhance the binding to platelets, with patients experiencing a range of clinical features depending on the resultant degree of instability of the domain. Mutations have been found to affect residues throughout AIM-A1, with varying effects causing the loss of autoinhibition of the A1 domain provided by the AIMS; detailed experimental study was conducted by Legan et al 2023 (108).

1.5 ADAMTS13 domain structure

The start of the ADAMTS13 gene encodes for a signal peptide and propeptide at the N-terminus, however these are shortly cleaved. The functional protein consists of a multi-domain structure with the N-terminal domains (MDTCS) responsible for proteolysis through a series of sites for VWF binding and cleavage, whilst the C-terminal domains regulate protein latency. Structural features of the N-terminal domains have been characterised in partial structures previously (97, 109). Recently the C-terminal CUB1/2 structures have also been resolved, so structural features known so far have been highlighted with further predications of interactions

suggested (110). Investigation into a complete structure of ADAMTS13 in both the latent and active forms would be useful to examine residues previously not thought to be relevant for activity. Each ADAMTS13 domain and features of structural and functional interest are subsequently explored in this chapter with Table 1.1 detailing ADAMTS13 exosites (secondary binding sites remote from the active site) for VWF binding.

Table 1.1 ADAMTS13-VWF exosites. List of exosite residues involved in each ADAMTS13 domain interaction with VWF.

ADAMTS13 domain	Exosite residues involved
Metalloprotease	L198, H224, E225, H228, L232, H234, L252, P256, L274
Disintegrin-like	R349, L350, V352
Cysteine-Rich	G471, A472, A473, A474
Spacer	Y658, R659, R660, Y661, Y665

1.5.1 Metalloprotease domain.

The most N-terminal of the domains is the MP, functionally responsible for the cleavage of VWF. The MDTCS E225Q structure (PBD: 6QIG), was previously resolved to 2.8 angstrom and was critical in understanding of novel features of the domain (97). ADAMTS13 only contains a very short propeptide domain (dispensable *in vitro*) that has very little homology to other ADAMTS propeptides (26), as mentioned previously this has limited function. Interestingly, the conserved feature of other metalloproteases, the cysteine-switch, is not found in ADAMTS13 and therefore does not control catalytic activity. The HEXXHXXGXXH MP zinc-binding motif and Met249 (named the Met turn) directly below this (providing structural integrity) are present which are conserved amongst ADAMTS proteases (111, 112). As mentioned previously, aligning of the VWF scissile bond to the active site is ensured through accommodation of P1 residue into S1 site (the catalytic centre at E225) and P1' into the S1' pocket shaped by D252-P256 (9). The Zn²⁺ ion adjacent to the active site, is co-ordinated by the three histidine residues (H224,

H228 and H234) of the zinc-binding motif (found in other ADAMTS proteins) (26, 111) (Figure 1.4A).

There is a distinct collection of residues named the ‘gatekeeper triad’ formed of charged residues Arg193, Asp217 and Asp252 (Figure 1.4B). Their side chains form intimate interactions between Ca²⁺ binding loops 180-193 and 231-263; the Ca²⁺ loop 180-193 occludes the active site cleft and blocks the S1’ pocket (95, 97). Petri et al (2019) proposed that this experimental structure is the latent form of ADAMTS13 that will not accommodate VWF (97). Allosteric activation of ADAMTS13 is achieved by VWF, which causes a conformational change allowing ‘opening of the gatekeeper triad’ (97). As well as the previously mentioned Ca²⁺ binding loop, there are 2 further Ca²⁺ binding sites on the reverse of the MP domain, in relation to the active site. These form a double binding site (specifically residues D166, D173, D284, E83, G168, C281 co-ordinate the Ca²⁺ ions) involving an irregular loop (residues 276–305) that connects to the disintegrin (Dis) domain (97) (Figure 1.4C).

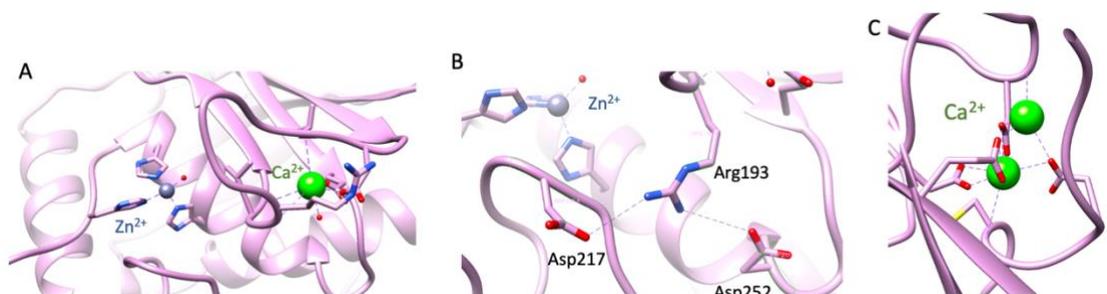


Figure 1.4 Molecular visualisation of ADAMTS13 metalloprotease domain features. ADAMTS13 Metalloprotease domain structure (PDB: 6QIG) there is a mutation at the active site of E225Q. (A) Active-site His residues (H224, H228 and H234) can be seen co-ordinating the Zn²⁺ ion, alongside the Ca²⁺ loop which occludes the active site in this latent formation (B) The gate keeper triad of residues Arg¹⁹³, Asp²¹⁷ and Asp²⁵² are responsible for the blocking of the active site cleft by the Ca²⁺ loop (C) The double Ca²⁺ binding loop is on the rear side of the MP domain.

1.5.2 Disintegrin domain

The ADAMTS13 Dis domain is atypical of others within the ADAMTS family, as many regions of homology are not present, such as the RGD tripeptide that is characteristic of disintegrins (113). Some highly conserved sequences are still

retained, with a similar overall structure of the Dis domain suggesting ADAMTS13 has evolved to function differently to other members of the family (111). It appears the Dis domain plays a role of functional importance both with enhancement of substrate binding (alongside cysteine-rich and spacer domains) as well as in the conformational activation of the MP domain and positioning of VWF into the active site to promote cleavage (96, 114).

The role of Dis domain in enhancing substrate binding also provides a level of protection against aberrant activation of ADAMTS13. Activation requires a series of complimentary interactions between specific residues on the VWF A2 domain with ADAMTS13 residues spread throughout the protein (109). The final exosite to be engaged is R349 (within the Dis domain) that interacts with Asp1614 in the VWF A2 domain (109). This interaction is supported by the approximate 26Å distance both between Arg349 and the Zn²⁺ ion in the catalytic cleft of ADAMTS13 and between Asp1614 and the VWF scissile bond (Tyr1605 and Met1606) (111, 114). This interaction induces the shift from the latent to the active form of the protein, with the structural rearranging of the MP domain including the disruption of the gatekeeper triad and rearranging of the Ca²⁺ binding loop (97, 115). The functionality of MP domain is also increased through the Dis domain, as the catalytic activity is enhanced with the domains in complex. Support was also provided by functional characterisation of exosites on the surface of ADAMTS13, a single point mutation R349D resulted in an almost complete loss of proteolytic activity(115).

1.5.3 Cysteine-Rich domain

The cysteine-rich (Cys) domain of ADAMTS13 contains 10 cysteine residues which are all paired to form disulphide bonds (109). This domain is generally homologous with others in the ADAMTS family, except within a functional region containing a non-conserved region (V-loop) harbouring the exosite residues G471-A472-A473-V474 which form a hydrophobic pocket to favour interactions with cryptic residues in the VWF A2 domain (I1642, W1644, I1649, L1650 and I1651) (116). The RGD loop mentioned previously, that is usually found in Dis domains, is present on the opposite side of the molecule to the V-loop in the Cys domain so is not functionally important in proteolysis(117).

The V-loop can be affected by mutations which may alter the functionality of VWF recognition and binding; for example, P475S substitution shows diminished substrate affinity and subsequent decreased activity (118). Mutagenesis studies have been carried out identifying the domain as important for VWF cleavage, but functional information for the importance of all residues is yet to be elucidated (118, 119). As well as affecting activity, incorrect folding or alteration to the 3D structure of ADAMTS13 leads to issues with secretion of protein indicated through further mutagenesis work (117).

1.5.4 Spacer Domain

The spacer domain is of functional importance for the interaction with VWF. Truncation of ADAMTS13 to MDTC domains leads to significantly reduced secretion and subsequent reduced VWF proteolysis, likely caused by incorrect protein folding and instability (117, 120). The MDTC crystal structure presented in Petri et al (2019) suggests the spacer exosite contains a cluster of hydrophobic residues on the beta-loops of the domain which mediates the binding of the VWF A2 domain to ADAMTS13 (97). Mutations affecting the spacer exosite can lead to a reduction in proteolytic activity, likely due to disruption of the interaction with the VWF exosites (121). If R660 is mutated (with a loss of positive charge) proteolytic activity is abolished suggesting this residue, and therefore the spacer exosite, is essential for activity through recognition of VWF (122).

Interestingly specific mutations of this spacer exosite can also lead to increased proteolysis of VWF (including substitutions R568K and F592Y), this is due to the strengthened hydrophobic interactions between the spacer domain exosite and the α 5-helix residues in VWF-A2 domain. The potential use of gain-of-function, GOF, ADAMTS13 variants as a treatment of TTP has been elucidated, specifically in patients with the acquired form of the disease, due to the resistance of the GOF variants to autoantibody targeting (121, 122). It could also be hypothesised that variants resistant to autoantibody targeting could also be useful for congenital TTP. To provide a viable treatment option, caution would be required regarding dosage to

ensure protein activity matched normal levels, rather than overactivation, to maintain the delicate balance that is haemostasis (121).

Differences in cleavage rates of MDTCS and MDTC (N-terminal domains without spacer domain) suggest there is an important role of spacer domain in catalytic efficiency (115). Under normal conditions ADAMTS13 has an inbuilt mechanism to maintain ‘global’ latency and prevent activation until required. This is provided by the curled interaction of the C-terminal CUB domains with the N-terminal spacer domain (Figure 1.5A). More detail on this interaction and mutations of the residues will be explored later in the chapter.

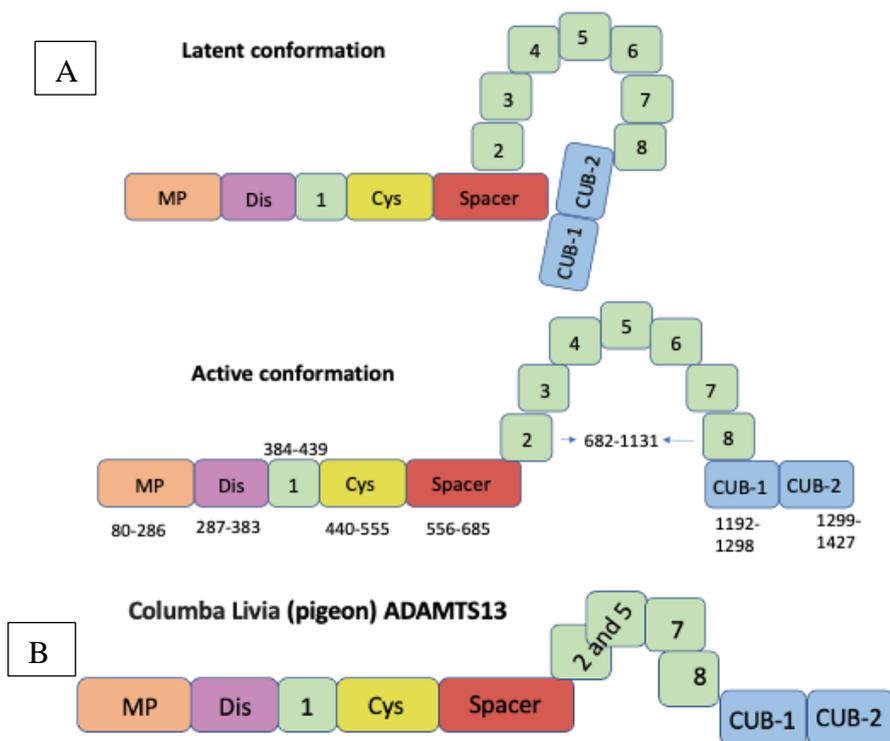


Figure 1.5 Schematic of the domain architecture of ADAMTS13. (A) Latent ADAMTS13 conformation involves the curled head-tail interaction between the Spacer-CUB domains and is not accessible to VWF, until the active form is achieved where the spacer-CUB interaction is broken. (B) The pigeon (Columba Livia) ADAMTS13 domain structure is illustrated as a comparison to the human form of the protein, with reduced number of TSP-repeats.

1.5.5 Thrombospondin type-1 repeat domains.

TSPs are composed of conserved tryptophan and cysteine residues in 2 anti-parallel strands (123). ADAMTS13 contains 8 TSPs, the first within the N-terminal domains between the Disintegrin-like and Cysteine-rich domains. TSPs 2-8 join the spacer domain to the C-terminal CUB-domains, with 3 linker regions between 2/3, 4/5 and 8/CUB-1 (124, 125). The TSP domains contain seven potential O-fucosylation sites that are critical for secretion of ADAMTS13, alongside N-glycosylation of multiple domains (18). The TSP linker regions are suggested to be important for flexibility of this region of the protein. (126)

TSP-1 is thought to function in facilitating ligand interaction. The domain contains of a WXXW motif; mutation of W387 in the motif significantly reduces ADAMTS13 secretion and reduces the binding affinity of ADAMTS13 to VWF, with resultant reduced proteolytic activity against FRETTS-VWF73 (20). It was later revealed this was a site of C-mannosylation which may explain the reduced secretion (for which C-mannosylation has been associated with) (20, 127). Features of C-terminal TSPs 2-7 are less characterised, as experimental structures have not been resolved. However, it is now understood TSP3-6 are dispensable, with Zhu et al reporting the minimal functional form of ADAMTS13 contains only 4 TSPs, closely resembling the pigeon ADAMTS13 gene (Figure 1.5B) (12). The role of these regions seems to be largely for flexibility and supporting the CUB-Spacer interaction which maintains latency of ADAMTS13 (128). Investigation of a minimal ADAMTS13 construct could be adopted into research where the human forms of the protein containing the large flexible regions are more difficult to characterise. Although T3-6 are dispensable, it can be predicted that mutations causing disruption of this flexible loop, allowing the CUB-spacer interaction, may be the cause of decreased activity through instability of the protein (12). Alternatively, some mutations may lead to reduced secretion due to abnormal folding of the protein. Mutation at R1060W is one of the most observed in TTP, and results in reduced protein secretion (12). Thus far no experimental crystal structures of these ADAMTS13 TSP repeats (2-8) have been resolved. The nature of this flexible region and its role in latency through accommodation of the Spacer-CUB interaction requires further investigation to both structural and functional features.

1.5.6 CUB Domains

There are two C1r/C1s, Uegf sea urchin fibropellins, and bone morphogenic protein 1 (CUB) domains at the C-terminal of ADAMTS13 (129). The CUB domains are specific to ADAMTS13, within the ADAMTS family, and previously less characterised than other domains until a recent paper by Kim et al 2021 (110). The CUB domains are involved in the initial binding of ADAMTS13 and the VWF D4CK domain, this preliminary binding results in conformational changes, occurring in both proteins, accommodating linear VWF to the active site of ADAMTS13 through disruption of the Spacer-CUB interaction (124). Previous work has revealed that ADAMTS13 folds its distal CUB domains to the proximal spacer domain to limit exposure of the N-terminal exosites until required for interaction with VWF(125).

Mutations in both CUB1 and 2 appear to affect protein secretion rather than protease activity directly supported by *in vitro* mutagenesis work based on TTP mutations (130, 131). Previously mentioned MDTCS constructs which lack C-terminal CUB domains still retained proteolytic activity supporting the role of CUB domains in latency rather than proteolysis (117). Loss of these C-terminal domains resulted in increased proteolysis under shear stress conditions, compared to WT ADAMTS13, due inability to control latency of the protein (normally achieved by the CUB-Spacer interaction) (132).

1.6 ADAMTS13 activation.

1.6.1 ADAMTS13 allosteric activation by VWF interactions.

The relationship between ADAMTS13 and VWF is notably unique, due to the intricate mechanism needed to activate both proteins. ADAMTS13 circulates in a latent closed formation and allosteric activation is achieved by a series of complementary exosite interactions with activated VWF, which therefore acts as both a substrate and an allosteric activator of ADAMTS13 (124, 125). Exosite interactions of ADAMTS13 and VWF have been researched previously using *in vitro* and molecular modelling techniques. Binding of CUB and D4-CK domain region initiates progressive unfolding, then exposes exosites in the A2 domain of VWF that engage with complementary sites on Spacer, Cys and Dis domains of

ADAMTS13 in that order (124, 133). It's predicted the spacer domain may recognise the VWF A2 domain when only partially unravelled towards its linear form and has the strongest exosite interaction with the A2 domain. The spacer and cysteine rich domain are in close proximity to each other with the hydrophobic pocket in the Cys domain next to be engaged by VWF. Finally, engagement of the Dis domain involves hydrophobic and ionic interactions with VWF before engagement of VWF at the MP active site.

Crawley et al. (2011) suggest the ADAMTS13 residues that are involved with each exosite interaction, these predicted exosite interaction residues are detailed in Table 1.1 and visualised in Figure 1.3. Further support has also been provided by more recent research, with these residues also included from Zander et al 2015 and Petri et al 2019. Although the VWF residues for interaction with ADAMTS13 have been mapped clearly to the VWF-73 peptide region, experimental evidence for the ADAMTS13 residues remains only a prediction based on modelling and *in vitro* experiments with no crystal structures of the complex to visualise interaction location and types available.

1.6.2 Latency of ADAMTS13

The latent state of ADAMTS13 is maintained by two levels of structural control; the global control of the Spacer-CUB domain interactions, and the local control within the metalloprotease domain composed of the gatekeeper triad interactions.

The Spacer-CUB interaction hides the multiple discontinuous exosites that are required for binding with VWF and the subsequent proteolysis. An initial interaction is needed between the D4-CK domain and CUB domains to release the Spacer-CUB interaction and initiate structural conformation changes of both proteins is required for activation (124). The residues involved in this interaction have been proposed by Kim et al (2021) from functional and docking data (Figure 1.6). This research suggests the interface consists of salt bridge interactions (E634 and D635 in the Spacer domain with K1252 and R1272 in CUB1), central hydrophobic interactions (L591 and F592, L637 and L668 in the Spacer domain and W1245, L1248, and W1250 in CUB1) as well as ionic interactions between a spacer domain loop and a

pocket in CUB-2 (R660-Y665 loop with R1326, R1361, E1387, and E1389 forming the pocket) (110). It's important to note these interactions are based on a model, and an experimental structure would be beneficial to visualise this curled head-tail interaction of ADAMTS13.

The second mechanism of latency is located within the MP domain, where the gatekeeper triad of residues restricts access of VWF to the active site until all exosites are engaged. The salt-bridge interactions between residues Arg193, Asp217 and Asp252 appear to stabilise the closure of the active site cleft until the adjacent Dis exosite is engaged allowing structural rearrangement of the MP domain to accommodate VWF in the active site for cleavage (96, 111). The precise mechanism of this structural rearrangement has not been elucidated but must involve the movement of the loop 180-193 away from the active-site cleft (Figure 1.4). The involvement of the Dis domain exosite in this activation is supported by experimental data showing that the MP and Dis domain in complex have greater proteolytic activity compared to the MP domain alone (115). The MP domain, and features mentioned here, have been resolved in the latent form but is expected to have a markedly different structure when active, as alluded to for accommodation of VWF. It was hypothesised by Halkidis et al, that the gatekeeper triad interaction could be an effect of the Fab used for crystallisation as well as a mechanism of latency by ADAMTS13 (134). Further investigation into this is therefore essential, with a need for a structure of the active form of ADAMTS13. This could be achieved utilising VWF to add stability through complex formation, negating the need for the stabilising Fab.

These protective mechanisms of latency dramatically decrease the chance of off-target or inappropriate cleavage of VWF or any other circulating proteins (125). It is also thought these latency mechanisms limit access of protease inhibitors to inactivate ADAMTS13 (there are no known inhibitors under normal circumstances) thus leading to the comparably high half-life of ADAMTS13 protein (135).

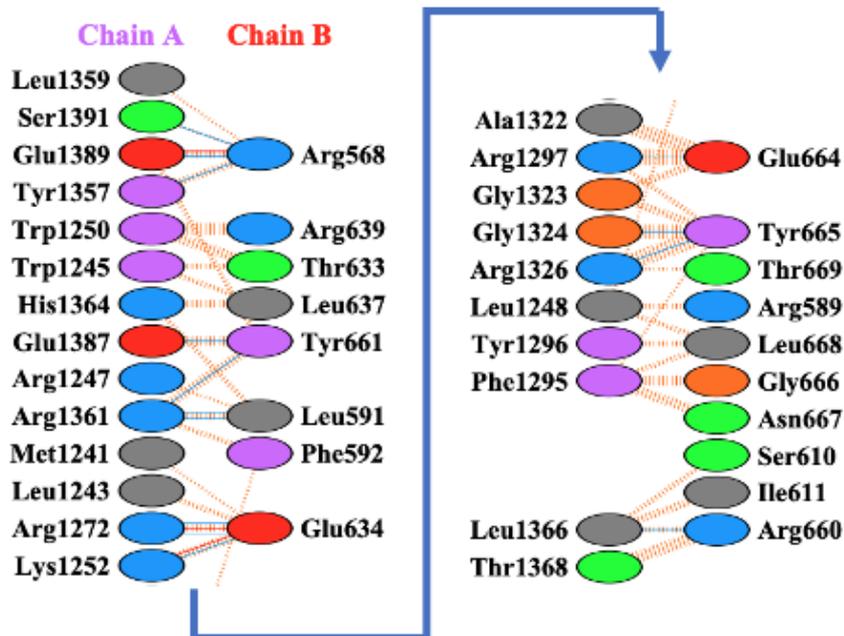
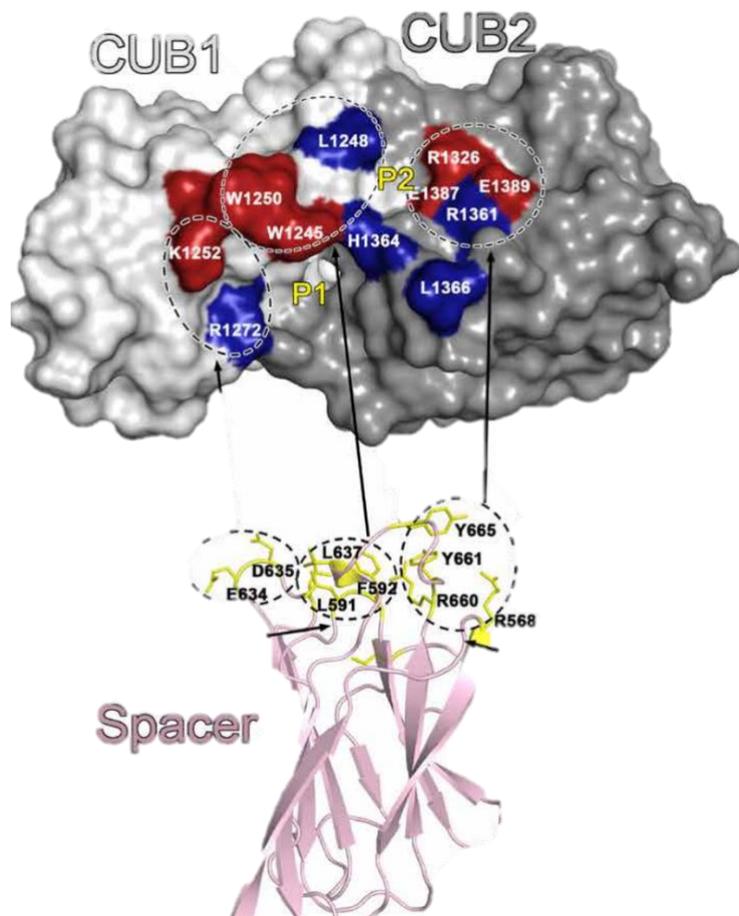


Figure 1.6 (A) The model adapted from Kim et al 2021, shows the proposed model of binding between the ADAMTS13 Spacer and CUB domains (PDB codes 6QIG and

7B01). This model generated using the ClusPro docking stimulation, suggests binding between both the P1 and P2 pockets on the CUB domains and the flexible loops of the Spacer domains. (B) A Ligplot schematic of interactions between ADAMTS13 CUB1/2 domains (chain A) and ADAMTS13 spacer (Chain B). Interactions are represented as lines with salt-bridges in red, hydrogen bonds in blue and non-bonded contacts as orange dashed lines. Figure adapted from Kim et al 2021(49).

1.7 Structure prediction and crystallisation

In recent years huge improvements have been made in structure prediction software which utilise a combination of functional data, genetic information and homologous sequences or structures to aid in accurate structure prediction. As well as this, methods for experimental determination with improvements in structural biology, instrumentation and software for analysis (allowing techniques including serial crystallography and room-temperature data acquisition) allow dynamic observations of protein structure with more relevance physiologically compared to other techniques (136, 137).

1.7.1 AlphaFold for structure prediction

An area of structural biology with ground-breaking improvements in software development is protein structure prediction. In 2020 the AlphaFold2 database was released revolutionising the capabilities of structure prediction. AlphaFold utilises the amino acid sequence of proteins to predict the 3D structure through physical and biological knowledge and sequence alignment. This AI network achieves accuracy not seen before in previous prediction software (138). Subsequently further upgrades have been made to include AlphaFold multimer enabling prediction of complex protein structures (139, 140). The AlphaFold predictions have accompanying scores including pLDDT which gives an estimate of local accuracy of prediction, with residue specific confidence intervals (pLDDT > 90 is very high confidence and pLDDT > 70 with moderate-high confidence) (138). A recent analysis of AlphaFold structures compared to experimental results suggest the accuracy has improved compared to previous predictions and can provide a hypothesis for experiments but

addressing of further uncertainties are required if they are to match experimental structure solution (141).

1.7.2 Previous crystallisation attempts of VWF and ADAMTS13

The VWF protein has 50 experimental structures deposited in the PDB to date ranging from single domain sections to entire tubule formations. The A1 domain has been deposited alone as well as in complex with GPIIb α and nanobody caplacizumab, whilst the A2 domain has structures deposited focusing on mutations present (Example PDBs: 1AUQ, 1M10, 7PNF). The A2 domain has no active open structures available and would likely need to be in complex with ADAMTS13 to achieve this stable linear conformation sufficient for structure resolution.

Crystallisation of ADAMTS13 domains has been previously successful, producing structures of DTCS, MDTCS and CUB domains, but no structure has combined the N-terminal and C-terminal domains thus far (PDB codes: 3VN4, 7B01, 3GHM, 3GHN and 6GIQ). Issues reported with crystallisation attempts of the complete ADAMTS13 protein are due to flexibility of many regions including the TSP repeats, as well as the MP domain that required a Fab fragment for previous crystallisation. It has been suggested the use of this Fab may produce a conformation not seen in the active or latent form of the protein so achieving a structure without this Fab would be beneficial to confirm physiological relevance of experimental structures (142). The extensive glycosylation network present on ADAMTS13 may also lead to crystallisation hindrance due to the flexible nature of these carbohydrate elements. The intrinsically disordered nature of ADAMTS13 means experimental data is the best way to confirm predictions of protein interactions with VWF are accurate and to investigate the conformational sensitivity achieved by allosteric activation of VWF (134, 143).

1.8 Aims of thesis:

I hypothesise the VWF protein will exhibit a novel conformation following regulation by ADAMTS13 or a nanobody. ADAMTS13 will also experience conformational changes upon interaction with VWF exhibiting a novel active conformation.

The aim of this study was to investigate the structure of both ADAMTS13-VWF and VWF A1-ND6 complexes. This study will design novel protein constructs of VWF utilising AlphaFold predictions and mutagenesis experiments to improve on attempts. These constructs will then be characterised and developed through protein expression and purification studies, as well as binding analysis for development of a stable VWF-ADAMTS13 complex suitable for crystallisation trials. Furthermore, this study will aim to crystalize VWF with a novel nanobody regulator (ND6) to reveal the complex crystal structure and subsequently interrogate structural relevant features of VWF. Investigating the regulation of VWF utilising structural biology techniques will provide an insight into how these proteins function in normal and disease states to inform potential treatment avenues.

2. Methods:

2.1 Analysis of ADAMTS13 variant data

2.1.1 Analysis of ADAMTS13 variants in the Ensembl database

The ADAMTS13 (ID: ENSG00000160323) variant information was obtained from the ensemble database (144). The variant table for ‘ADAMTS13’ available at (https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000160323;r=9:133414358-133459402) was subsequently filtered to include only entries with evidence from ‘phenotype or disease’ and ‘literature’ to ensure variants were confirmed through sequencing and/or literature as well as an additional filter to ignore any entries with ‘blank amino acid co-ordinates’. This data was downloaded, replicate information was then removed (by ensuring all variant IDs were unique) which resulted in 2674 entries. This data was then analysed for themes in the information on gene variation type, consequence on protein, domain locations of mutation and REVEL class. Simple statistical analysis was conducted to compare frequency of features shared between mutations, and graphs were then generated from the results.

2.1.2 Analysis of ADAMTS13 mutations in the Ensembl database

To focus on specific patient mutations in TTP, a stricter search to the previous ADAMTS13 variant analysis was conducted, this time filtering for ‘likely pathogenic’ mutations but also importantly with available ‘citation’ information. This allowed cross-referencing and provided more detailed information on patient features such as disease onset or severity, as well as protein structure or function information, with protein activity or secretion results available. This information was collated into a summary table and themes of mutation domain location, disease onset and severity, and structural defects were investigated in an attempt to highlight any associations between genotype and phenotype.

2.2 Designing constructs for VWF-ADAMTS13 interaction studies.

2.2.1 Using AlphaFold for prediction of ADAMTS13 and VWF structures.

Proteins can be searched in the AlphaFold database based on their Uniprot ID number, and an AlphaFold predicted structure can be downloaded alongside

confidence scores, such as for ADAMTS13 (145). However, where structures were not available, such as for new constructs being developed, the ColabFold platform was used for AlphaFold v1.5.2 software (138, 146). This required only the input of the protein amino acid sequence and produces 3 top models with associated pLDDT and PAE scores, indicative of AlphaFold confidence (138). Structure prediction was run for DTCS-CUB ADAMTS13 protein and VWF-2GKG proteins. The sequences used for construct generation can be found in Supplementary Table 1. The AlphaFold Multimer (v.2.3.0) docking software utilises the same ColabFold software for running AlphaFold and requires the input of both protein amino acid sequences to generate the three-dimensional models. Each AlphaFold prediction run was performed without templates, with automatic parameters and the default number of cycles to generate five relaxed predictions. Docking was conducted for 2GKG-VWF constructs with ADAMTS13 MDTCS, MP-Dis and Cys-Spacer domains.

2.2.2 Guided docking of ADAMTS13 and VWF

Three-dimensional models of the interaction between ADAMTS13 (MDTCS domains PDB: 6QIG) and VWF A2 domain (73 amino acids in a simple linear conformation) were calculated utilising the ROSIE docking server (147). A combined PDB file of the two proteins with a ‘reasonable guess for the starting position’ is required for uploading onto the sever. This was achieved manually by manipulation of proteins in Pymol due to the ease of manipulating the linear VWF, but with more complex proteins would require an initial round of docking utilising other software (as indicated in the documentation) (147, 148). The output file was then interrogated in Pymol. Ligplot was utilised via the EMBL-EBI PDBsum server with upload of the ROSIE output file to obtain interaction graphics (149).

2.2.3 Sequencing and amplifying DNA constructs.

6xHis-SUMO-VWF-73 construct was obtained from a collaborator at Imperial College London (Crawley), produced previously by cloning the 73 amino acid sequence of the human VWF A2 domain fragment (D1596-R1668) into pET-SUMO plasmid (Invitrogen) containing an N-terminal 6xHis-SUMO tag. This was sent for sequencing using primers T7F (TAATACGACTCACTATAGGG) and T7R

(GCTAGTTATTGCTCAGCGG). Once the DNA sequence was verified, amplification of DNA was performed using *E.coli*.

Heat transformation of *E. coli* was conducted as follows. Competent Novablue *E.coli* cells were thawed on ice from -80 stocks. Petri dishes were made with Lysogeny Broth (LB)-media/agarose mixture and 0.05mg/ml kanamycin. 2µl of DNA stock (100ng/µl) was added to an aliquot of 100µl Novablue cells and incubated on ice for 10 minutes. This was then heat shocked by heating at 42°C for 45 seconds and then returned to ice. Cells were then allowed to recover for 1 hour at 37 °C in LB. The transformed cells were then spread in a thin film onto the prepared petri dishes and left overnight at 37°C. Single colonies were used to inoculate 10ml liquid LB broth with 0.05mg/ml kanamycin added and incubated overnight at 37°C while shaking at 180rpm.

For sequencing, a miniprep kit from Gen-elute (Sigma-Aldrich) was used, and the standard procedure followed. 5ml cells were harvested by centrifugation at 12,000 x g and the pellet resuspended in 200µl buffer. 200µl lysis solution was then added and the solution gently mixed. Cell debris was precipitated by adding a neutralisation buffer producing a viscous white precipitate. The spin column was prepared by washing with buffer, the cell solution added and centrifuged at 12,000 x g for 1 minute. Flow through was discarded, an ethanol solution used to wash the column and then the DNA was eluted using 50µl dH₂O and sent for sequencing or stored immediately at -20 °C. To produce DNA stocks at larger volumes, Genelute midi kit was also used (requiring 100ml of culture) to give a larger volume of DNA compared to the miniprep kit.

2.2.4 Site-directed mutagenesis of 6xHis-SUMO-VWF-73

Plasmid vectors containing wild type templates of 6xHis-SUMO-VWF-73 were used to introduce mutation/s in their DNA using an Agilent site-directed mutagenesis kit (Aligent QuikChange II Site-Directed Mutagenesis Kit). Primers were designed using Aligent primer design tool and TakaraBio primer design (Table 2.1) for constructs consisting of mutations Y1605C, YM1605-6CC and G1629E (Schematic in Figure 2.1 and full sequence in Supplementary Table 1). Following the protocol

provided: 10-50ng of template DNA was mixed with 10x reaction buffer, 0.5 μ M primer each (forward and reverse primer), dNTPs (0.2 mM), 1 μ l of PfuUltra HF DNA polymerase (2.5 U/ μ l) and centrifuge briefly prior to running the PCR protocol (Table 2.2). The amplified PCR products were Dpn1 treated for 1 hour at 37 °C to get rid of the methylated parent DNA template, leaving the amplified unmethylated PCR mutant product. PCR amplification was followed by transformation into Novablue cells as described previously.

Table 2.1 Mutagenesis primers used for generation of mutant 6xHis-SUMO-VWF-73 constructs. For double mutants, PCR protocol was followed twice following successful sequencing of the first mutation (Y1605C or YM1605CC then G1629E).

Mutation	Primers (5'-3')
Y1605C	5'-ttccggtgaccatgcagaccaggttgggc-3' 5'-gccaacctggtctgcatggtcaccggaa-3'
YM1605-6CC	5'-tggctctgctgcgtcaccggaaatcctgcctctg-3' 5'-tgacgcagcagaccaggttgggcgcctg-3'
G1629E	5'-cattagggcccacttcaatgggcaccacc-3' 5'-ggtggtgccattgaagtgggccctaatg-3'

Table 2.2 - PCR protocol for mutagenesis reactions using QuikChange II Site-Directed Mutagenesis Kit.

Segment	Cycles	Temperature	Time
1	1	95	30 sec
2	12 (point mutation)	95	30 sec
	16 (single AA change)	55	1 min
	18 (multiple AA insertions or deletions)	68	8 mins

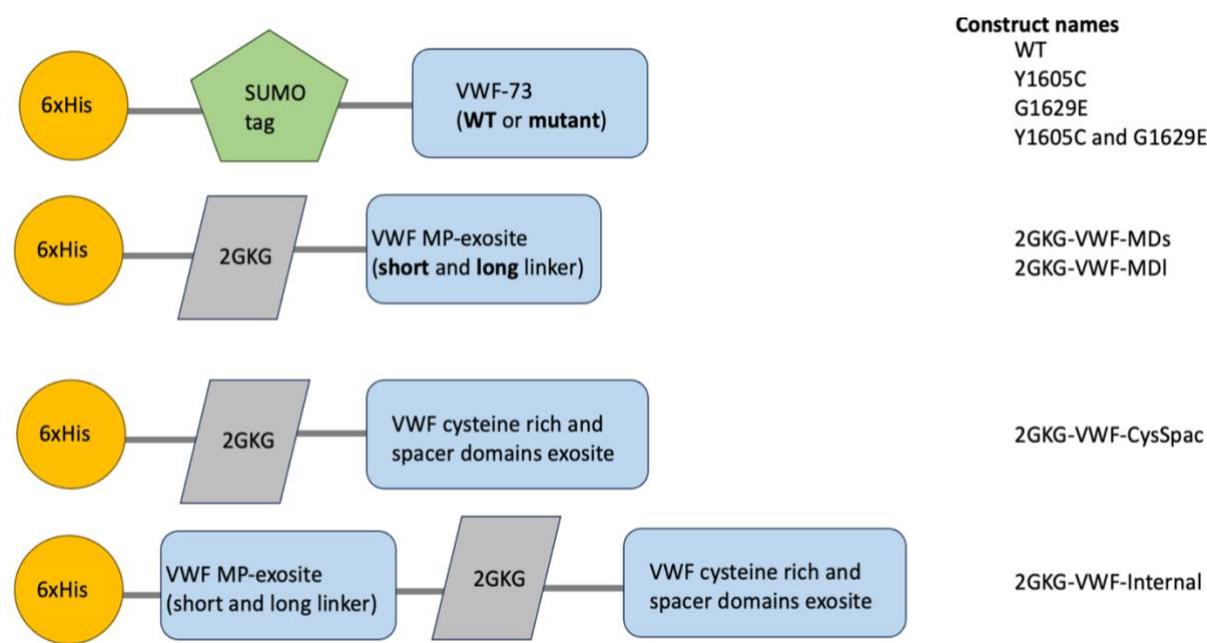


Figure 2.1 Schematic representation of VWF A2 constructs. Each section of the construct represented by a different shape with each having its own use either as a tag (6xHis, SUMO and 2GKG) or the functional VWF element (shown in light blue).

2.2.5 VWF-cloning 2GKG.

The 2GKG tag (named after the PDB entry) is formed of the receiver domain of *Myxococcus xanthus* social motility protein FrzS and is ~14kDa in size. Previous studies have identified this tag as an ideal candidate to improve protein expression and crystallisation (150). The VWF-2GKG construct insert sequences can be found in Supplementary Table 1 and a schematic in Figure 2.1. The cloning experiments were carried out in collaboration with a lab member. The VWF-2GKG inserts and pet-25b(+) were both set up for a double digestion reaction with NheI/XhoI in NEBuffer 2 + rAlbumin at 37 °C. The insert was amplified with a PCR reaction and then purified using QIAQuick PCR purification kit. The vector was purified using the QIAQuick gel extraction kit. A ligation was carried out utilising the T4 ligase following the NEB protocol and the resultant reaction was transformed into NovaBlue cells. Sequencing with T7F and T7R primers was used to confirm resultant colonies contain the insert.

2.3 Protein expression

2.3.1 *E.Coli* transformation and glycerol stock generation

The 6xHis-SUMO-VWF73 (WT and mutants) and 2GKG-VWF constructs were transformed into BL21 (DE3) *E.coli* cells (Invitrogen) using heat transformation as described in Chapter 2.2.3. The transformed cells were spread onto LB agar plates with 50µg/ml kanamycin, or ampicillin 100ug/ml for 2GKG constructs, to select for successfully transformed cells. Transformant colonies were obtained by incubation overnight at 37°C. Single colonies were inoculated into 10ml LB, with respective antibiotics and cultured overnight at 37°C with shaking at 180rpm. 1ml of overnight cultures was centrifuged (10min, 13,000g) and resuspended in 300µl 30% glycerol/LB solution and were then frozen at -80°C to generate a glycerol stock of the transformed expression strain.

2.3.2 *E.Coli* protein overexpression

Glycerol stocks were scraped and used to generate an overnight culture. This was added to 3x flasks of 800ml of LB with 50µg/ml kanamycin or ampicillin 100ug/ml and allowed to grow until OD600 of 0.6 was reached at which point 1mM IPTG was added, and cells were grown at 22°C for ~20hours. Cells were harvested by centrifuging at 4600 x g, 30mins, 7°C. Media was discarded and the cell pellet was resuspended into falcon tubes for further centrifugation at 5000 x g 30mins, 7°C and the supernatant discarded. The pellet could then be stored at -20°C until required for purification.

2.3.3 *Drosophila* S2 protein expression system (transfection and overexpression)

ADAMTS13 constructs are expressed in the *Drosophila* S2 protein expression system (Invitrogen). The MDTCS domains (G78-P682) cloned into the insect cell expression vector pMT-BiP-PURO, which enables secretion of the protein, was obtained from a collaborator (Crawley, Imperial college London). This construct contains the active site mutation E225Q to prevent cleavage of the substrate but enable complex formation. They also provided the human full-length (FL) WT ADAMTS13 protein (151). The DTCS-CUB construct (sequence in Supplementary Table 1) was designed and provided by a lab member.

2.3.4 Calcium phosphate transfection into S2 cells

The constructs were amplified using the miniprep kit as described previously with ampicillin as the selective agent and sterilised using ethanol precipitation, so they were suitable for cell transformation.

Stock *Drosophila* S2 cells were taken from a liquid nitrogen cell bank, then thawed to inoculate 5ml CM media (Schneider's complete medium (Gibco) containing L-glutamine (15mM)). This was then incubated at 28°C for 40 min. DMSO was removed by centrifuging at 1000 x g for 2 mins, the supernatant discarded, and the pellet resuspended in fresh CM media. 10µl is then taken from the culture, mixed with 10µl trypan blue (ThermoFisher) on a slide and analysed using an automated cell counter (TC10, BioRad). The cell culture was incubated at 28°C until an approximate cell count of 1×10^7 cells per ml of culture (maintained through splitting 1:2) and cells were abandoned when percentage of live cells dropped below 70%. Cells were also observed via light microscope to check for infections.

DNA transfection into insect cells was achieved following a calcium phosphate protocol from Invitrogen. To form a calcium-phosphate DNA precipitate, 36µl 2M CaCl₂ was added to ~ 20ug DNA, this was made up to 300µl total with purified water. This solution was then mixed dropwise with 2X HEPES buffered saline with continuous mixing. The solution was incubated at room temperature for 30-40 minutes, mixed gently and added dropwise to 3-4ml of cell culture in log phase ($2-4 \times 10^6$ cells/ml). After incubating over the weekend at 28°C, the CaCl₂ was removed by centrifuging the cell culture at 1000 x g for 1-2 min, and then resuspending in fresh S2 media. This was repeated twice to ensure all CaCl₂ was removed.

Cells were then maintained at a density of approximately at 1×10^7 cells, with CM replaced with Express Five™ serum-free medium (Gibco, ThermoFisher) with the addition of L-Glutamine (200 mM, 100 mL), penicillin (1000 units) and streptomycin (10 mg) per litre of media. Following a few passages of cells with the new media, the cells were grown for overexpression.

2.3.5 Drosophila Protein overexpression

The culture was grown with addition of Express Five medium as described above until the desired total volume was achieved (>2 L) at a density of $2-4 \times 10^6$ cells/mL. Induction was achieved by adding copper sulphate to a final concentration 500 μ M and flasks were incubated (28°C, 180 rpm) for 7 days.

The Drosophila S2 system secretes the expressed protein directly into the media. Media was centrifuged (4600 xg, 1hr, 7°C) to separate supernatant and cells, the supernatant was then decanted and centrifuged again (4600 xg, 30 mins, 7°C) prior to dilution 1:1 with His-Excel Buffer (20mM HEPES, 100mM NaCl, 2mM CaCl₂). The solution was left to stand for 30 minutes to allow any precipitate to form and then filtered using Corning® 1000 mL Vacuum Filter/Storage Bottle System (0.22 μ m Pore). ADAMTS13 WT FL constructs were not purified so media was centrifuged and filtered as above and then frozen in aliquots of 500 μ l and 1ml, as well as stock volumes of 50ml.

2.4 Protein purification

2.4.1 ADAMTS13 MDTCS E225Q construct

A HisTrap Excel column (Cytivia) was pre-equilibrated with 5 CV of His-Excel Buffer A (20mM HEPES, 100mM NaCl, 2mM CaCl₂). The diluted and filtered protein-containing media was loaded onto the equilibrated column using a peristaltic pump. The column was then transferred to an AKTA system, and a wash was performed with Buffer A + 20mM imidazole. Elution was performed step wise with a wash at 150mM, elution at 250mM and wash at 500mM imidazole utilising His-Excel Buffer B (addition of 0.5M imidazole). Protein-containing fractions from the Ni-affinity purification were combined and spin-concentrated to reduce volume. The sample was then applied to Superdex S200 16/600 column (Cytivia) (Buffer contained 20mM HEPES 150mM NaCl, 2% glycerol) at 0.5ml/min. The protein containing fractions eluted after approximately 70mL. Presence of MDTCS E225Q was confirmed by SDS-PAGE gel and samples were concentrated to ~4mg/ml and flash frozen in liquid nitrogen for later use.

2.4.2 VWF-73 constructs

Both 6xHis-SUMO-VWF73 and 2GKG-VWF constructs were purified following the same method. Harvested bacterial cell pellets were resuspended and lysed by sonication at $\lambda = 15\mu\text{m}$ with 20 second on/off pulses for 7 minutes. Cell debris was pelleted by centrifugation at 14000rpm for 1 hour at 7°C. The supernatant was passed through a 0.22 μm filter before being loaded onto a 5ml HisTrap Excel column (Cytivia) which was equilibrated with 20mM Tris-HCl pH 8 and 200mM NaCl (Buffer A). Fractions were then eluted from the column using a 0-100% Buffer B (addition of 500mM imidazole) gradient run over 100ml on an AKTA system. Fractions containing high peaks of UV were selected and analysed on an SDS-PAGE gel.

Fractions containing the desired protein were selected and further purified by gel filtration on HiLoad[®] 16/600 Superdex[®] 75 pg (S75) column (Cytivia). The S75 column was equilibrated with 1.5 Column Volumes of buffer (20mM Tris, 150mM NaCl and 1% glycerol). Protein-containing fractions were loaded onto the column and then eluted using the buffer previously mentioned. Fractions of high UV were selected and analysed on SDS-PAGE gel, and fractions containing VWF constructs were concentrated down and stored at -80°C until required.

2.4.3 ADAMTS13 DTCS-CUB

The DTCS-CUB construct was transiently expressed, and a small-scale expression and purification was performed utilising a HisTrap Excel column (Cytivia) and Superdex S200 16/600 column. Following these columns, a band containing the suspected DTCS-CUB construct was excised and sent for Mass Spectrometry analysis (details below in Chapter 2.5.1).

2.5 Complementary techniques for construct characterisation

2.5.1 Mass Spectrometry

SDS-PAGE gels were run with purification products from WT 6xHis-SUMO-VWF-73, MDTCS E225Q and DTCS-CUB concentrated size exclusion purification fractions. Bands were excised from the gel and send to University of York

Metabolomics & Proteomics Lab for analysis by MALDI-TOF mass spectrometry. Analysis carried out at York checked results against provided sequences for constructs and where no match was found checked against the UniProt database (152).

2.5.2 Western blot

MDTCS E225Q, WT MDTCS and all VWF constructs were subjected to western blotting analysis to confirm the correct protein was expressed and purified.

An SDS-PAGE gel was prepared and run, as before, until the stage of staining and visualisation. The SDS-PAGE gel was instead transferred to nitrocellulose membrane 0.45µm pore size using the Trans-Blot Turbo (Bio-Rad). The equipment (including cathode and anode stacks) was assembled according to manufacturer's instructions using Tourbin buffer and a 13-minute transfer method was used. The membrane was removed and washed for 10mins in TBST and subsequently submerged in Blocking Buffer (2% milk) and incubated for 2h RTP. After removal of the blocking buffer, the membrane was covered with the primary antibody, 6x-His Tag Polyclonal antibody (Invitrogen #PA1-983B) and incubated with gentle agitation ON at 4 degrees. The membrane was washed thoroughly for 3x 10-minute intervals with TBST, prior to incubation with the secondary antibody, Goat Anti-Mouse IgG(H+L) (ThermoFisher #31430), (dilution 1 : 5000) for 40 minutes with gentle agitation. The membrane was washed 3x 10-minute intervals with TBST and 1x 10 minutes with TBS. The visualisation kit (SuperSignal™ West Pico PLUS Chemiluminescent Substrate, ThermoFisher) was then used as per manufacturer's instructions and visualised using CanoScan LiDE 210 (Canon).

2.5.3 Expression and purification of SUMO-protease ULP-1

The expression and purification of the SUMO protease largely followed the protocol used in Lau et al 2018 with the plasmid pCDB327 gifted from Christopher Bahl (Addgene plasmid # 113671 (153)). The Nickel excel column was utilised instead of Ni-NTA beads and a gradient elution performed using a 0-100% Buffer B (addition of 500mM imidazole) over 100ml with 2ml fractions collected. The S75 column was ran as mentioned and aliquots frozen for use in cleavage assays.

2.5.4 SUMO cleavage of 6xHis-SUMO-VWF73 constructs

The ULP-1 SUMO protease was utilised for cleavage of the SUMO tag from the VWF constructs. SUMO protease was added to 6xHis-SUMO-VWF73 constructs as well as a control without VWF, at the recommended ratio of 1 unit enzyme to 2µg fusion protein. Samples were incubated in SUMO protease buffer (500 mM Tris-HCl pH 8.0, 2% NP-40, 1.5 M NaCl and 10 mM DTT) for 3 hours at RTP following Invitrogen protocol for ULP-1 cleavage (Cat #12588018).

2.5.5 WT ADAMTS13 cleavage assay on VWF

To establish cleavage of VWF by ADAMTS13 a cleavage assay was established for qualitative analysis of the presence of cleavage. Incubation of 5µg of VWF constructs for 0h and ON with 10µl of CM containing active ADAMTS13 was set up alongside a control of VWF constructs alone. 15µl aliquots were taken at each time point and the reaction was ended by the addition of EDTA. All samples were then analysed using an SDS page gel to visualise the presence of a cleavage product and reduction in uncleaved VWF.

2.5.6 ITC and SPR

Isothermal calorimetry was conducted on an PEAQ-ITC in a 200 µL reaction cell loaded with 200µl MDTCS ΔE225Q at 25 °C. The titration was performed using 19 injections of 6xHis-SUMO-VWF73 constructs in the same buffer (20mM Tris-HCl pH7.5, 150mM NaCl) at 150s interval. Binding isotherms were analysed using the Microcal PEAQ-ITC software and fit to a single site binding model.

Surface Plasmon Resonance spectroscopy experiments were performed and analysed by a colleague in the lab, but briefly the experimental method involved the following. SPR spectroscopy was performed utilising a BIAcore 2000 (Cytivia) using CM5 sensor chips (Cytivia). Ligand (MDTCS E225Q) was amine-coupled to the CM5 surface using Amine coupling kit (Cytivia). Amine coupling occurs between NHS-esters and amine side chains allowing immobilisation of the ligand to the surface. The analyte (6xHis-SUMO-VWF constructs) was prepared in the SPR running buffer. The flowrate for analyte injections was between 40-50ul/min. Regeneration utilised Glycine pH2.0. Sensorgrams were analysed and fitted using BIAevaluation software on double reference subtracted curves.

2.5.7 Complexing for size exclusion analysis

VWF 6xHis-SUMO constructs and MDTCS E225Q were mixed and subsequently incubated at RTP for 4 hours at ratios of 1:1, 2:1. The samples were then loaded onto both Superdex™ 75 10/300 GL column (Cytivia) and Superose™ 6 Increase 5/150 GL Columns (Cytivia) for size exclusion analysis.

2.5.8 Crystallisation trials

MDTCS (E225Q) at 2mg/ml was incubated with WT and Y1605C mutant 6xHis-SUMO-VWF73 (4mg/ml) for 4 hours at room temperature for complex formation. This was then concentrated down to 5mg/ml for use in crystallisation experiments, the results of ITC experiments were also concentrated and utilised. The mosquito robot (TTP Labtech) was used to dispense volumes of 0.3µl of the VWF-MDTCS complex and 0.3µl of crystallisation solution in a 96-well plate MRC plate (Molecular Dimensions) utilising sitting drop vapour diffusion method. The crystallisation screens used included JSCG+ and Pact Premier (Molecular Dimensions) and were stored at 15 and 20°C.

2.6 Techniques for crystallisation and characterisation of VWF-ND6 complex

2.6.1 Crystal growth and X-ray crystallography data collection

Two VWF-nanobody complexes (VWF AIM-A1/ND4 and VWF AIM-A1/ND6) were provided by R. Li (Emory) for structural determination by X-ray crystallisation studies. Sequences found in Supplementary Table 1.

The AIM-A1/ND6 complex was concentrated to ~16 mg/ml for crystallization trials using commercial crystallisation screens (JCSG+, Morpheus III, Proplex, Midas+) from Hampton Research (Aliso Viejo, CA) and Molecular Dimensions (Sheffield, UK) in sitting-drop crystallization plates at 15°C. Multiple fine shard-like crystals grew from condition G9 Morpheus III. Consisting of 1.2 % Cholic acid derivatives mix (3% w/v CHAPS, 3% w/v CHAPSO, 3% w/v Sodium glycocholate hydrate, 3% w/v Taurocholic acid sodium salt hydrate), 0.1 M Buffer System 3 pH 8.5 (Imidazole; MES monohydrate) and 30 % Precipitant Mix 1 (40% v/v PEG 500* MME; 20 % w/v PEG 20000). Crystals were harvested, then flash-frozen in liquid nitrogen for data collection on beamline i24 at Diamond Light Source. Due to the

long thin nature of the crystals, a line scan was utilised to collect data along the length of the crystal. Diffraction data were collected from multiple crystals to 3.4Å resolution.

The AIM-A1/ND4 complex was concentrated to ~8mg/ml for crystallization trials using commercial crystallisation screens (JCSG+, Morpheus III, Proplex, Midas+) from Hampton Research (Aliso Viejo, CA) and Molecular Dimensions (Sheffield, UK) in sitting-drop crystallization plates at 15°C. Multiple crystals grew from condition G12 Stura Footprint Eco screen (0.2 M Ammonium sulphate 36 % w/v PEG 4000). Crystals were harvested and 30% glycerol added as a cryo-protectant, then flash-frozen in liquid nitrogen for data collection on beamline I04 at Diamond Light Source.

The VWF AIM-A1 protein was also concentrated to 3 mg/ml and set up in crystallisation trials using commercial screens (JCSG+ and PactPremier) from Molecular Dimensions (Sheffield, UK) in sitting-drop crystallization plates at 18°C. Crystal clusters were observed in condition JCSG+ E6 (0.2M Zinc acetate dihydrate, 0.1 M Imidazole 8.0, 20 % w/v PEG 3000). An optimisation experiment was set up with conditions from 0.12- 0.22M zinc acetate hydrate and 8-22% PEG-3000 with 0.1 M Imidazole 8.0.

2.6.2 Generation of model for structure solution.

Models for both ND4 and ND6 nanobodies were generated using AlphaFold by providing the sequence for the nanobody (138). Models were then subsequently trimmed for use in structure solution using Sculptor in PHENIX to remove a terminal flexible region (154).

2.6.3 Data processing and model building with X-ray crystallography data.

Datasets collected from the diffracted A1-ND6 crystal were indexed and phased automatically using autoProc and STARANISO was used to correct anisotropy, the output of these processes is a .mtz file that can be used in later steps. Indexing and phasing are now often done automatically with powerful tools utilised automatically available at the beamline. Briefly, indexing involves giving the crystal orientation

and an idea of unit cell dimensions as well as crystallographic symmetry and index which will be subsequently refined. Anisotropy refers to the directionally dependant properties of the data and should be corrected for before continuing data processing (155)(156). For phasing using molecular replacement, Phaser was initially given a VWF A1 model (PDB: 5BV8 (157)) and an AlphaFold model of ND6 as well as sequences for both proteins and the .mtz file. 2 copies of each protein were found in the unit cell. This was then followed by initial model building in COOT (158) and refinement using both `real_space_refine` in PHENIX (159) and REFMAC5.0 (160). For the addition of O-glycans; A2G and SIA monomers were imported, and LINK files were generated in COOT using ACEDRG (161) with the restraint files then used in subsequent refinement. The A1-ND6 structure was then validated using MOLPROBITY in PHENIX. Table 6.1 gives information on data statistics, for space group P 2 21 21, The R/R-free remained around 0.3 (0.45 outer shell), with a low completeness (68%).

Images were subsequently reindexed and phased to give the space group P 1 21 1 following processing with AutoProc and STARANISO. The statistics for this data are presented in Table 6.1, notably the beta angle here is 89.95° compared to 90° allowing processing in an alternative space group. The final model from the previous processing was utilised as a search model in molecular replacement in the new unit cell and space group, 4 copies of the A1-ND6 complex were found. The initial R/R-free factors improved compared to previous processing, so this new space group was adopted for subsequent processing.

The completeness and multiplicity for a single crystal were suboptimal so further data sets were analysed in the space group P1 21 1 from another crystal (data was indexed and phased using xia2 automatically in ISPYB). In CCP4 data was then reindexed into the space group P 1 21 1 using the data from the previous crystal with space group P1 21 1 as a reference mtz file. Molecular replacement using MOLREP with the search model from the previous partial solution was utilised. This was followed by LORESTR refinement and manual model building in COOT before refinement cycles with PDB-REDO and then TLS. Subsequently, extra regions of density were observed between nanobody chains, the PDB file was analysed using

the CheckMyBlob software which suggested cholic acid with low confidence (25%) due to breaks in density (162). However, as the result matches with the crystallisation buffer utilised (cholic acid derivatives mix), CHAPS (CPS) and cholic acid (CHD) were both imported and modelled into the density manually in COOT. Regions of density could also be seen around T1468 and S1263, known O-glycosylation points of AIM-A1, so the monomer A2G was imported and manually modelling into the density using COOT. The valency of these monomers was set to 0 as surrounding density is not well defined. Final refinement was run again using LORESTR refinement followed by final manual adjustments in COOT before final refinement cycles with TLS. The A1-ND6 structure was then validated using MOLPROBITY in PHENIX and wwPDB Validation Service. Pymol and Chimera were used to visualise structures and create figures (163). The complex structure will subsequently be deposited into the PDB.

For the A1-ND4 complex, diffraction data collected indicated the harvested crystals contained only ND4 nanobody alone without VWF A1 due to the small unit cell, so was processed subsequently for this nanobody alone. Diffraction data were collected from multiple crystals and processed with Xiaai, CCP4 suite to 1.15 Å resolution. The structure was then solved using molecular replacement (Phaser) with the nanobody model generated using AlphaFold. Followed by manual model building using COOT and refinement with REFMAC. The ND4 structure was then validated using MOLPROBITY in PHENIX (Table 6.1)

2.6.4 Aligning for comparison of complex structures.

Aligning of structures for visual comparison was performed in UCSF Chimera utilising the matchmaker tool using default settings (163). The find Hbonds function was utilised to visualise bonds between residues.

2.6.5 Analysis of interactions

Interactions between AIM-A1 and ND6 as well as ND6 and cholic acid were interrogated utilising both EMBL-EBI PDBSUM and PDBePISA servers with PDB files uploaded of the complexes and detailed interaction information between residues as the output (149, 164).

3. Linking genotype and phenotype: exploring the TTP mutations, and the structural effect on ADAMTS13.

3.1 Introduction

TTP is a rare blood disorder caused by reduced or abolished ADAMTS13 function. The diagnosis of TTP largely relies on ADAMTS13 activity (threshold set at <10%), due to lack of uniformity in symptoms and respective severity across patients. Symptoms often may be misdiagnosed as multiple thrombotic microangiopathies (36). In patients presenting with thrombocytopenia and MAHA, TTP should be suspected, and subsequent diagnostic laboratory testing carried out (165). Initially this should involve an ADAMTS13 activity assay, but consequently also an ADAMTS13 functional inhibitor assay or anti-ADAMTS13 antibody assay to investigate the origin of the disease to differentiate between the congenital or acquired forms (166). If the antibody assay comes back as negative, genetic testing should be carried out to gain the sequence of ADAMTS13 and check for ADAMTS13 mutations. Congenital TTP should also be expected if there is a family history, neonatal onset or pregnancy induced disease which are all associated factors (36). Since 2009, the UK TTP registry has been used to record clinical information from each TTP patient with aims of determining incidence and epidemiology of TTP, as well as looking for association in ADAMTS13 mutations and clinical features amongst other aims (167).

The range in severity and symptoms associated with TTP is extremely broad with mild symptoms of bruising to fatal manifestations of stroke or heart attack. The initial presentation of symptoms often includes neurological issues including headaches and confusion, as well as general weakness and abdominal pain, bruising and bleeding are also common (35, 168). Currently, there is no clear link between the genetic defects of TTP mutations and the disease course the patient will then experience. Due to the rare nature of this disease novel mutations are reported each year, with no associations thus far found in clinical data or research to inform how the disease is likely to progress. Recent research to understand more about the detrimental effects of mutations on the ADAMTS13 protein have been categorised into either reducing protein secretion or activity (131, 169). The structural effects behind what causes the difference is however less clear, as mutations affecting

secretion have been found across all domains for ADAMTS13 (131, 170-177). The function of the N-terminal MP domain (and by close association the Dis domain) is the metalloprotease activity to cleave VWF, consequently mutations can often affect protease activity (95). Conversely, mutations in the CUB domain have highlighted the importance of these domains for stability of protein, which in turn affects secretion (178, 179). There remain large gaps of information around ADAMTS13 mutations and the subsequent effect on protein secretion and activity.

Even with the heterogeneity reported in many papers about disease onset, clinical manifestation and number of relapses, the treatment course for TTP patients is largely universal. This commonly consists of fresh frozen plasma to replenish functional ADAMTS13, as well as intermediate purity factor VIII concentrate which is used in the UK as some patients do not respond sufficiently to fresh frozen plasma (FFP) (180, 181). Immune-modulating drugs and corticosteroids are used for treating acquired TTP due to the autoimmune response experienced in these patients and has further evidence for prophylactic use (182, 183). Whilst some patients remain in long-term recovery from this treatment pathway, others experience multiple periods of relapse. With limited alternative options for treatment, they then require prophylactic plasma exchange. This treatment course (varying frequencies of prophylactic FFP from multiple times a week to once a month) has reported decreased risk of severe clinical manifestations of the disease such as ischemic stroke in some patients, but limited success in others (38). The time-consuming invasive treatment route for patients with congenital TTP therefore needs radical updating to help improve patient quality of life.

Research into novel TTP treatments is ongoing, and will also be further explored in Chapter 6, with multiple avenues open for exploration. Novel treatments hope to appeal to a larger number of cTTP patients and provide a longer-term cure without recurrence of relapse due to the functional ADAMTS13 protein in the blood. The idea of gene therapy for treating cTTP was reviewed recently by Dekimpe et al (2022) in which the advantages and challenges of this approach were summarised; as well as detailing the studies to date on gene therapy and murine models for ADAMTS13 replacement (184). Briefly, current issues include longevity of ADAMTS13 expression, as well as a desirable delivery system — which are

common challenges faced for gene therapy. Results thus far however have shown this therapy would be an attractive alternative to the current treatments, with more research needed into recombinant ADAMTS13 in pre-clinical analysis and confirming the best ADAMTS13 construct for expression. The other favourable therapy option being explored is replacement of functional ADAMTS13 in the form of human recombinant ADAMTS13 protein. Clinical trial results for multiple phases have been published on the drug ADZYNMA, with the FDA highlighting the improvements in prophylactic use compared to plasma-based therapies for cTTP patients over 12 years old (72). Whilst ADZYNMA/rADAMTS13 is now approved for use by the FDA, it is yet to receive approval for use in the UK (73). Furthermore, the exact mechanism by which ADAMTS13 and VWF interact and therefore the precise mechanism of the interaction between the rADAMTS13 and VWF is currently still unknown. This highlights a pivotal area of research to better understand this new treatment option, with potential for improvements and personalisation in the future, allowing for the heterogeneity in TTP treatment response.

To help establish the background for understanding the landscape of TTP, the following chapter will interrogate and summarise clinical data from TTP patients, including patient mutations, symptoms, and disease course. This will help inform what further evidence is needed to better understand how the ADAMTS13 protein is affected by the TTP mutations and utilising this information to progress disease treatment.

3.2 Results

3.2.1 Summary of themes in ADAMTS13 mutations in the Ensembl database.

To provide an overview on the known ADAMTS13 mutations, the Ensembl database was utilised to collect information on ADAMTS13 mutations with evidence such as citation information or inclusion in another database such as dbSNP from NCBI or HGMD-PUBLIC (185-187). To summarise, there were 2674 entries that met the above criteria and were used for these simple statistics, with multiple entries included if reported in multiple patients. The results of the analysis on ADAMTS13 mutation entries are presented in graphical form (focused on gene variation type,

consequence on protein, domain locations of mutation and REVEL class) in Figures 3.1 and 3.2. Whilst often used interchangeably, in this thesis a variant is defined as a permanent structural alteration in DNA and whilst a mutation refers to the same event but at an amino acid level.

From these analyses, it is clear there is a wide spread of both mutation types and domain location, with disease causing effects seen across the ADAMTS13 protein. As expected from previous research, mutations are found across all domains, with MDTCS domains at the N-terminus and the C-terminal domains TSP2-8 and CUB1/2 each with a spread of the mutations across these domains. (Figure 3.1A). Furthermore, pathogenic changes can be instigated from single amino acid substitutions to larger frameshift mutations or insertions and deletions. Whilst missense mutations are the largest in number, there were also a large proportion of synonymous variants reported that would not be expected to cause TTP as the coded amino acid would be maintained (Figure 3.1B).

The largest proportion of variant type was single nucleotide polymorphism (SNP), followed by somatic SNV (single nucleotide variant) (Figure 3.1C). By definition, SNPs are found in at least 1% of the population which can indicate that some variants in ADAMTS13 are tolerated and do not cause TTP. To better understand the likelihood of mutations resulting in TTP, the mutation type was investigated as well as analysis of the predicted effect on pathogenicity (REVEL score) and effect on protein structure (SIFT score). The REVEL score classification of mutations supports the variant type information, with a larger proportion of variants predicted to be benign rather than disease causing (33% and 9% respectively) (Figure 3.2A). When looking at the effect of mutations on protein structure, the SIFT score classification of mutation indicated that a similar proportion of mutations are predicted to be deleterious or tolerated (20% and 21% respectively) (Figure 3.2B). Many of these variants are unclassified for pathogenicity or effect on protein structure (58% for both REVEL and SIFT scores) as supporting data and experimental results were lacking (Figure 3.2A-B). To look specifically at mutations that are known to cause TTP, a stricter search was used to include only cited pathogenic information to look at any correlation between genotype and phenotype.

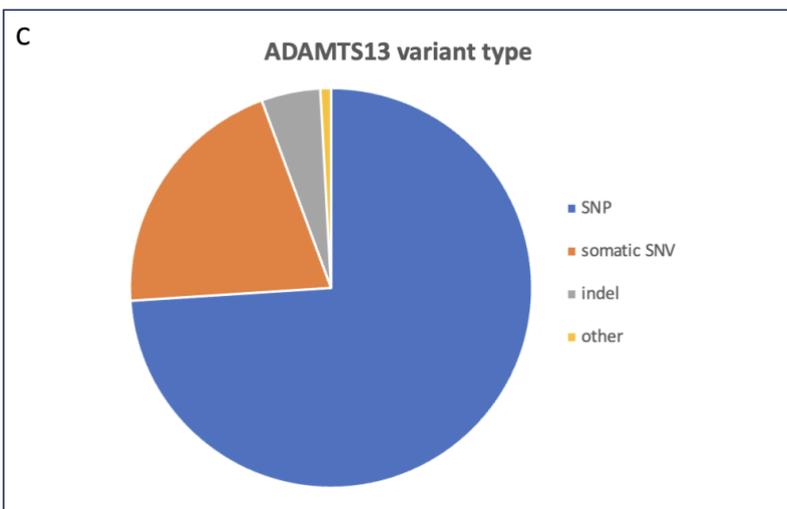
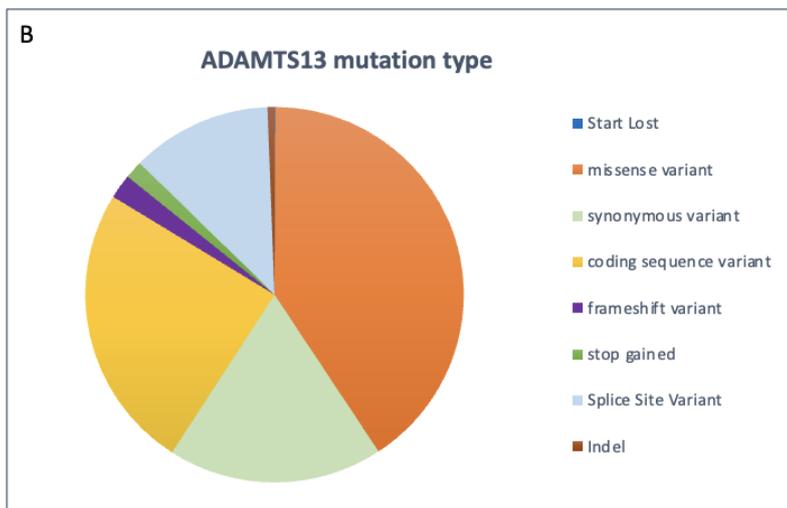
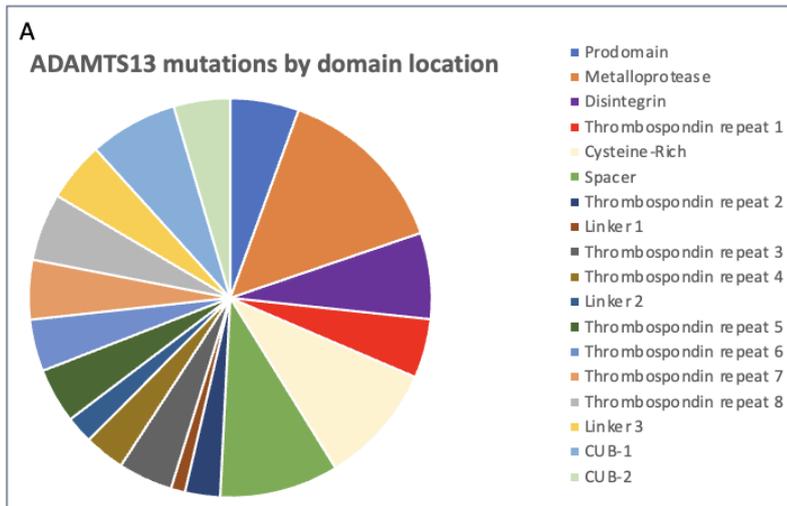


Figure 3.1 Graphical representation of ADAMTS13 mutations when analysed by (A) mutation domain location (B) mutation type and (C) variant type. Mutations

are spread through ADAMTS13 domains. The majority of ADAMTS13 mutations are missense, with a majority of the ADAMTS13 variants being SNPs.

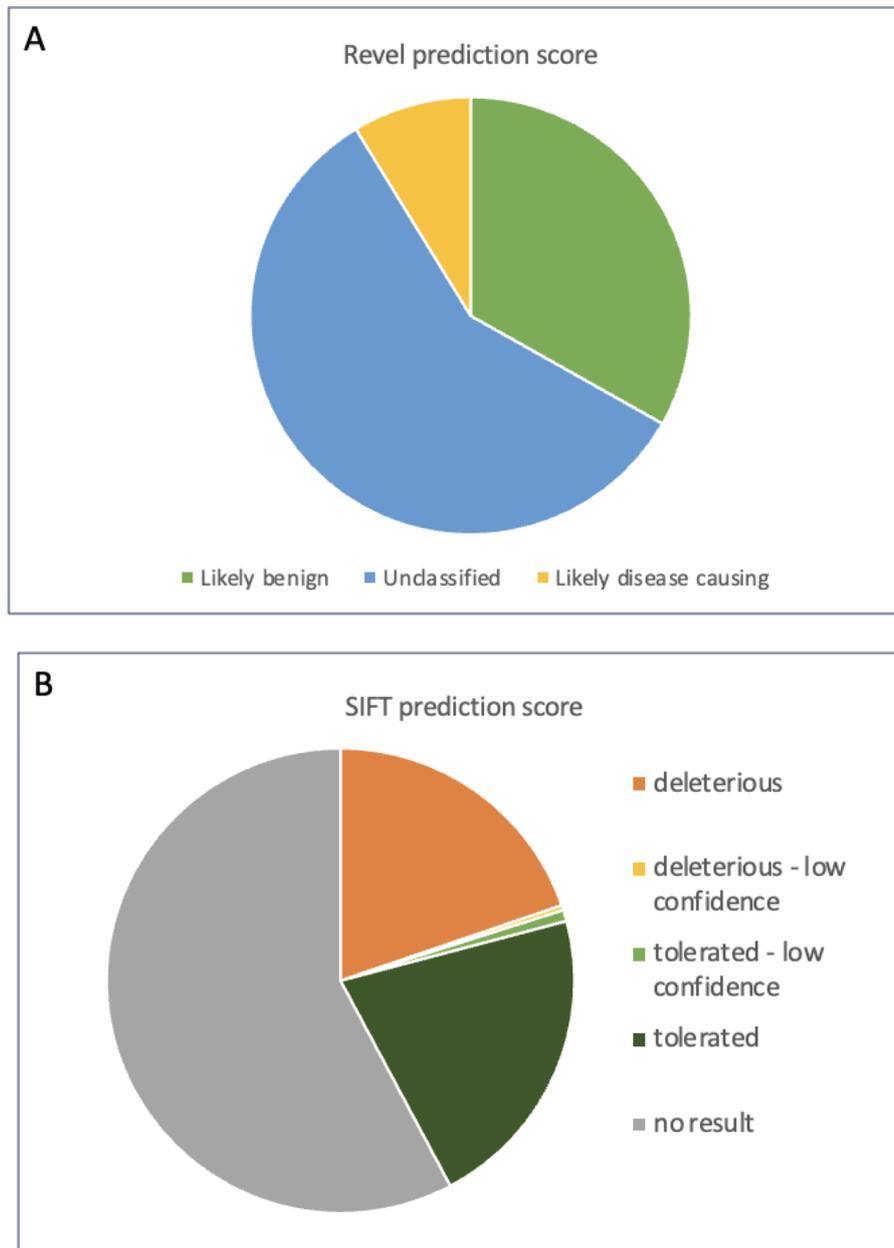


Figure 3.2 Graphical depiction of ADAMTS13 mutations when sorted by outcome of (A) REVEL prediction (B) SIFT prediction. The most common result was unclassified with no score given for both predictions, followed by likely benign for pathogenicity (REVEL score) and similar percentages of deleterious or tolerated structural changes (SIFT score).

3.2.2 Assessing the link between genotype and phenotype in TTP patients

To enable a better look at how specific mutations are affecting the ADAMTS13 protein, a summary table of mutations and their structural and clinical effects was collated. This builds on the table found in Markham-Lee Z. et al (2022) (174) and includes any mutations found in the literature with sufficient information available to allow analysis (Table 3.1). A visual representation of these mutations sorted by domain location can be found in Figure 3.3. Overall, there was limited reported mutation information available with detailed enough evidence, but some themes could be reported within and across domains. Numerous mutations affect the large MP domain (18 mutations reported in detail), with mutations affecting the Zn²⁺ binding site or the Ca²⁺ loops which are both crucial for proteolytic activity and thus function of ADAMTS13. The reported mutations in the TSP repeats, spacer and CUB domains are reported or predicted to interfere with protein folding and stability. The Cysteine-rich domain mutations also likely lead to issues with protein folding and stability as well as disruption of VWF recognition at the exosite.

Across all domains apart from TSP-1, there was evidence of both neonatal and adult-onset disease, so the domain location of the mutation cannot provide indication of disease onset at this stage. Patient information relating to ADAMTS13 activity and records of disease relapses are not consistently reported in literature, so it is therefore hard to draw any associations from these factors, but generally both can be seen in both adult and early onset disease. Neonatal onset disease was seen more frequently in all domains apart from CUB. There could therefore be an association between the CUB domains and a milder cause of disease, however more data is needed to confirm this. Further analysis needs to include more patient mutation entries and ensure there is not a false association due to the literature more frequently reporting the cases of severe, neonatal onset disease compared to patients with milder disease courses. The availability of *in vitro* data highlighted the higher frequency of secretion to activity disruption in this dataset, whereas routine activity assays are not able to differentiate between inactive or missing protein in samples.

The structural effect on ADAMTS13 from mutations was reported in the literature based on a variety of techniques from modelling the mutations onto the structure, to performing molecular dynamic simulations. This means direct comparisons can be

challenging to make; however, some associations can be formed from the available information. Commonly reported or predicted structural changes in ADAMTS13 included disruption of both intra/inter domain interactions, exosite residues and disulphide bond interactions. As expected, a loss of a large region or truncation of the protein (such as mutation Q1041*) did cause severe disease compared to mutations causing just a localised change in folding such as R1060W and C977F. The disruption of cysteine interactions was predicted in cases such as in R692C and R1123C potentially introducing a disulphide interaction, or C908Y where an interaction may be lost, resulting in severe disease with younger onset. In comparison mutations with only predicted localised misfolding such as R1219W, Q448E and R1060W generally had adult onset or milder disease. However, R268P and I673F resulted in abolished secretion *in vitro*, where the stability of the protein may be compromised and thus disrupts normal secretion resulting in neonatal onset disease. Mutations predicted to affect the active site such as (R102S and G236C) led to differing courses of disease severity, the current models of ADAMTS13 are in the closed state so it may yet be unclear which residues are involved in the active site interaction, both directly and indirectly, which may explain differing disease courses when understood.

Table 3.1: ADAMTS13 mutations in TTP patients. The selection of mutations was based on available phenotype data as well as cited evidence to give examples across domains. Where *in vitro* work has been carried out this has been included to help understand the effect of the specific mutation on the protein, especially in the case of heterozygous mutations.

Affected domain	DNA Change	Amino acid change	Homozygous or compound heterozygous	Patient characteristics	Predicted structural effect on the protein	Other supporting evidence	Functional studies?	References
MP	262G>A	V88M	Heterozygous	Pregnancy/infection induced; long-term remission achieved	Potential disruption of Ca ²⁺ loop	2 nd mutation was G1239V, reduced secretion and activity <i>in vitro</i>	Yes	Rurali et al 2015, Cataland SR et al 2014
	305G>A	R102H	Homozygous	Pregnancy induced	Disruption of MP-Dis interaction or indirectly disrupt active site	<i>In vitro</i> partially reduced secretion	Yes	Hing et al 2013 (188), Underwood 2015,
	304C>A	R102S	Homozygous	Adult onset, severe seizures, Undetectable ADAMTS13 activity and severely reduced ADAMTS13 antigen levels	Disruption of MP-Dis interaction or indirectly disrupt active site	Molecular dynamics experiments showing disruption of binding with Dis domain, severely reduced secretion and activity <i>in vitro</i>	Yes	Elbaz et al 2020, Da Waele et al 2022
	332G> A	G111E	Heterozygous	Neonatal onset, haemolytic anaemia and impaired renal function, multiple	Potential disruption of active site either directly or indirectly	Parent is heterozygous asymptomatic	No	Hou and Du 2020

				relapses, 6% ADAMTS13 activity		carrier of the mutation		
356C>T	S119F	Homozygous	Diagnosis at 17yo, severe complications	Potential disruption of active site either directly or indirectly	Reduced secretion with reduced catalytic efficiency	<i>Yes</i>	<i>Meyer et al 2008</i> <i>Hing et al 2013</i>	
414+1G>A	Splice variant	Homozygous	Neonatal onset <3% activity	Abolished splicing at the exon 4–intron 4 boundary	Improper splicing at exon/intron 4.	<i>No</i>	<i>Matsumoto et al 2004</i>	
428T>C	I143T	Homozygous	Recurring relapse	Potential disruption of active site either directly or indirectly	Intracellular retention and proteasome degradation	<i>Yes</i>	<i>Underwood 2015,</i> <i>Hing et al 2013</i>	
448T>C	S150P	Homozygous	Neonatal onset	Potential disruption of active site either directly or indirectly	NA	<i>No</i>	<i>Hing et al 2013</i>	
530A>G	Y177C	Heterozygous	Childhood onset, neurological symptoms, renal impairment, ADAMTS13 activity <5%	Disruption of Ca ²⁺ binding loop, destabilising the protein	Severely reduced secretion <i>in vitro</i> , molecular modelling suggests reduced intradomain interactions	<i>Yes</i>	<i>Wang et al 2023</i>	
577C>T	R193W	Homozygous	Pregnancy induced	Disruption of active site	Reduced secretion with no detectable activity	<i>Yes</i>	<i>Hing et al 2013,</i> <i>Matsumoto et al 2004</i>	

581G>T	G194V	Heterozygous	Childhood onset, infection induced and experienced relapse	Potential disruption of active site or Ca ²⁺ binding loop	Also reported in healthy controls so likely not the contributing mutation	No	Ma et al 2006
587C>T	T196I	Patient 1: Homozygous Patient 2: Heterozygous	P1: Infant diagnosis, P2: infant diagnosis and activity at <5%.	Disruption of S1 subsite and therefore VWF interaction	<25% activity reported <i>in vitro</i>	Yes	Levy et al 2001, Pimanda et al 2004, Camilleri 2012
649G>C	D217H	Heterozygous	infant onset, infection induced, severely reduced ADAMTS13 activity and undetectable antigen levels	Potential disruption of Ca ²⁺ binding loop	Reduced secretion and activity <i>in vitro</i>	Yes	Camilleri 2012
695T>A	L232Q	Homozygous	Neonatal onset	Disruption of active site either directly or indirectly	NA	No	Hing et al 2013
703G>C	D235H	Homozygous	Neonatal onset	Disruption of active site either directly or indirectly	NA	No	Hing et al 2013
706G>T	G236C	Homozygous	Pregnancy induced	Disruption of active site either directly or indirectly	NA	No	Hing et al 2013
749C>T	A250V	Heterozygous	Recurring relapses, activity <3%	Close proximity to Zn ²⁺ ion may compromise proteolytic activity	Markedly reduced activity	Yes	Uchida et al 2004, Shelat et al 2005
803G>C	R268P	Heterozygous	Neonatal onset and frequent relapses	Disrupted protein folding	Abolished secretion and reduced activity	Yes	Kokame et al 2002, Hommais et al 2007

Dis	932G>A	C311Y	Homozygous	Neonatal onset	Disulphide bond formation affected	NA	No	<i>Hing et al 2013, Assink et al 2003</i>
	1045C > T	R349C	Heterozygous	Pregnancy induced, activity <1%	Dis domain exosite is altered affecting VWF interaction	NA (but R349D has complete loss of activity)	No	<i>Fujimura et al 2009, Akiyama et al 2009</i>
	Not stated	C365del6	Heterozygous	Adult onset	Deletion of residues disrupts stability of Dis domain	Daughter heterozygous. Severely reduced secretion and abolished activity	Yes	<i>Tao et al 2006</i>
TSP-1	1177C > T	A393*	Heterozygous	Neonatal onset and severe disease course, activity <4%	Severely truncated protein	NA-hypothesised as inactive	No	<i>Hassenpflug et al 2018</i>
	1192C>T	R398C	Heterozygous	Neonatal onset, chronic Microangiopathic haemolytic anaemia	Introduction of cysteine	Abolished secretion	Yes	<i>Camilleri et al 2012 Plaimauer et al 2006</i>
	1225C>T	R409W	Homozygous	Neonatal onset, <5% activity	Hydrophobic substitution of residue may cause destabilisation	Reduced secretion	Yes	<i>Camilleri et al 2012</i>
	1308G>C	Q436H	Homozygous	Neonatal onset	May affect the local environment around disulphide bonds.	NA	No	<i>Hing et al 2013</i>
-	1331G>A	splice variant	Heterozygous	Acute episodes, <3% activity	Improper mRNA formation due to splice variation	NA	No	<i>Uchida et al 2004</i>
Cys	1342C>G	Q448E	Heterozygous	Patient 1: ischemic stroke	Introduction of acidic charged amino acid, in	Predicted to be tolerated as seen in general	Yes	<i>Camilleri et al., 2012</i>

				Patient 2 (son of patient 1): asymptomatic	a conserved region of the protein	population, only slightly reduced activity		
	1345C>T	Q449*	Homozygous	<3% activity in homozygous patient and 45-60% in heterozygous parents.	Loss of C-terminal domains	Normal secretion but very low activity detected, parents heterozygous	Yes	<i>Kokame et al 2002</i>
	1370C>T	P457L	Heterozygous	Neonatal onset, multiple relapses, activity <5%	Potential disruption of exosite or instability due to intradomain interactions	Parent carries the mutation, and had 50% ADAMTS13 activity	Yes	<i>Manea et al 2006, Assink et al 2003</i>
	1492C>T	R498C	Heterozygous	Neonatal onset, multiple relapses of severe thrombocytopenia and one episode with ischemic cerebrovascular accidents	Protein instability due to lacking intradomain interactions	Molecular dynamics simulations carried out indicated protein instability, severely reduced secretion <i>in vitro</i>	Yes	<i>Schelpe et al 2008</i>
	1523G>A	C508Y	Heterozygous	<3% activity	Affects proper protein folding	<i>In vitro</i> secretion is abolished	Yes	<i>Kokame et al 2004</i>
Spacer	1783-1784delTT	L595fs	homozygous	Very low activity (< 0.1 U/mL)	Truncated protein due to frameshift	NA	No	<i>Savasan et al 2003</i>

	1787C>T	A596V	Two unrelated families, one heterozygous and one homozygous	2 unrelated patients both with <5% activity and neonatal onset	A larger residue substitution may alter tertiary structure	Greatly reduced secretion and activity	Yes	<i>Veyradier et al 2004, Camilleri et al 2012</i>
	1973A>G	Y658C	Homozygous	Pregnancy induced, <5% activity	Cysteine may introduce disulphide bond formation	NA	No	<i>Ho Lee et al 2011</i>
	2017A>T	I673F	Heterozygous	Neonatal onset	Affects protein folding	Abolished secretion	Yes	<i>Matsumoto et al 2004</i>
TSP-repeats	Not stated	W688X	Homozygous	Neonatal onset, ADAMTS13 activity <1%	Prematurely truncated protein	Severely reduced activity under shear stress.	Yes	<i>Meyer et al 2008,</i>
	2068G>A	A690T	Heterozygous	Pregnancy induced, with multiple miscarriages		Severely reduced secretion <i>in vitro</i>	Yes	<i>Camilleri et al 2012</i>
	2074C>T	R692C	Homozygous	Neonatal onset, ADAMTS13 activity <1%	Disulphide bond formation may be affected	Severely reduced secretion	Yes	<i>Meyer et al 2008 and Hing et al 2013</i>
	2260T>C	C754R	Homozygous	2 patients (siblings), both with childhood onset and very low ADAMTS13 antigen levels	Loss of cysteine or introduction of larger side chain may affect local environment	Severely reduced secretion <i>in vitro</i>	Yes	<i>Camilleri et al 2012</i>
	Not stated	C804R	Homozygous	Neonatal onset, ADAMTS13 activity <1%	Disulphide bond formation maybe affected	Severely reduced secretion	Yes	<i>Meyer et al 2008</i>
	2723G>A	C908Y	Heterozygous	neonatal onset, <3% activity	Disruption of disulphide bond	Abolished secretion	Yes	<i>Matsumoto et al 2004</i>

	2851T>G	C951G	Heterozygous	Neonatal onset with chronic relapse, <7% activity	Loss of cysteine may affect local interactions or disruption of disulphide bridge	Severely reduced activity <i>in vitro</i>	<i>Yes</i>	<i>Shelat et al 2005, Motto et al 2002, Levy et al 2001</i>
	2930G>T	C977F	Homozygous	Pregnancy induced	Loss of cysteine may affect local interactions or disruption of disulphide bridge	Consanguineous parents heterozygous with subnormal ADAMTS13 activity, severely reduced secretion <i>in vitro</i>	<i>Yes</i>	<i>Camilleri et al 2012</i>
	2930–2935del	C977-R979 delinsW	Homozygous	<20% activity reported	Truncated protein	NA	No	<i>Peyvandi et al 2004</i>
	3033delC	C1012AfsX109	Homozygous	Neonatal onset, low ADAMTS13 antigen levels detected, Microangiopathic haemolytic anaemia	Loss of TSP-repeat may destabilise protein	Parents heterozygous asymptomatic carriers	No	<i>Rashid, Mushtaq and Masoori 2020</i>
	3121C>T	Q1041*	Heterozygous	Neonatal onset, haemolytic anaemia and impaired renal function, multiple relapses, 6% ADAMTS13 activity	Truncated protein	Parents heterozygous asymptomatic carrier of the mutation	<i>No</i>	<i>Hou and Du 2020</i>
	3178C>T	R1060W	Both types reported	Very low ADAMTS13 activity, asymptomatic until	Localized misfolding within ADAMTS13	Severe intracellular	<i>Yes</i>	<i>Savasan et al 2003, Scully et al 2014, Lotta et al 2010</i>

				pregnancy, patients have milder disease courses	due to change in charge	retention, <5% secretion		
	3367C>T	R1123C	Patient 1: Heterozygous Patient 2: Homozygous	Patient 1: Neonatal onset, <3% activity, recurrent episodes Patient 2: childhood onset, multiple relapses, chronic kidney disease, patient died	Creation of mixed disulphide bond	Abolished secretion	Yes	<i>Matsumoto et al 2004, Rurali et al 2015</i>
CUB-1	3655C>T	R1219W	2 patients, brothers, both homozygous	Both patients with adult onset, both suffered relapse before complete remission	Potential localized misfolding within ADAMTS13 due to change in charge	Abolished secretion	Yes	<i>Donadelli et al 2006, Rurali et al. 2015</i>
	3716G>T	G1239V	Homozygous	Infant onset, recurring relapse, activity was <1%	Affecting protein structure through folding	Severely reduced secretion and reduced activity	Yes	<i>Meyer et al 2008, Donadelli et al 2006</i>
	c.3923G>A	G1308D	Heterozygous	Pregnancy induced, <3% activity	Affects protein folding and stability	Severely reduced secretion	Yes	<i>Jiang et al 2020</i>
CUB-2	4143dupA	E1382Rfs	Both reported	Homozygous with neonatal onset and more severe clinical course than heterozygous	Truncated protein	Severely reduced secretion and activity	Yes	<i>Hassenpflug et al 2018 and Peyvandi et al 2004</i>

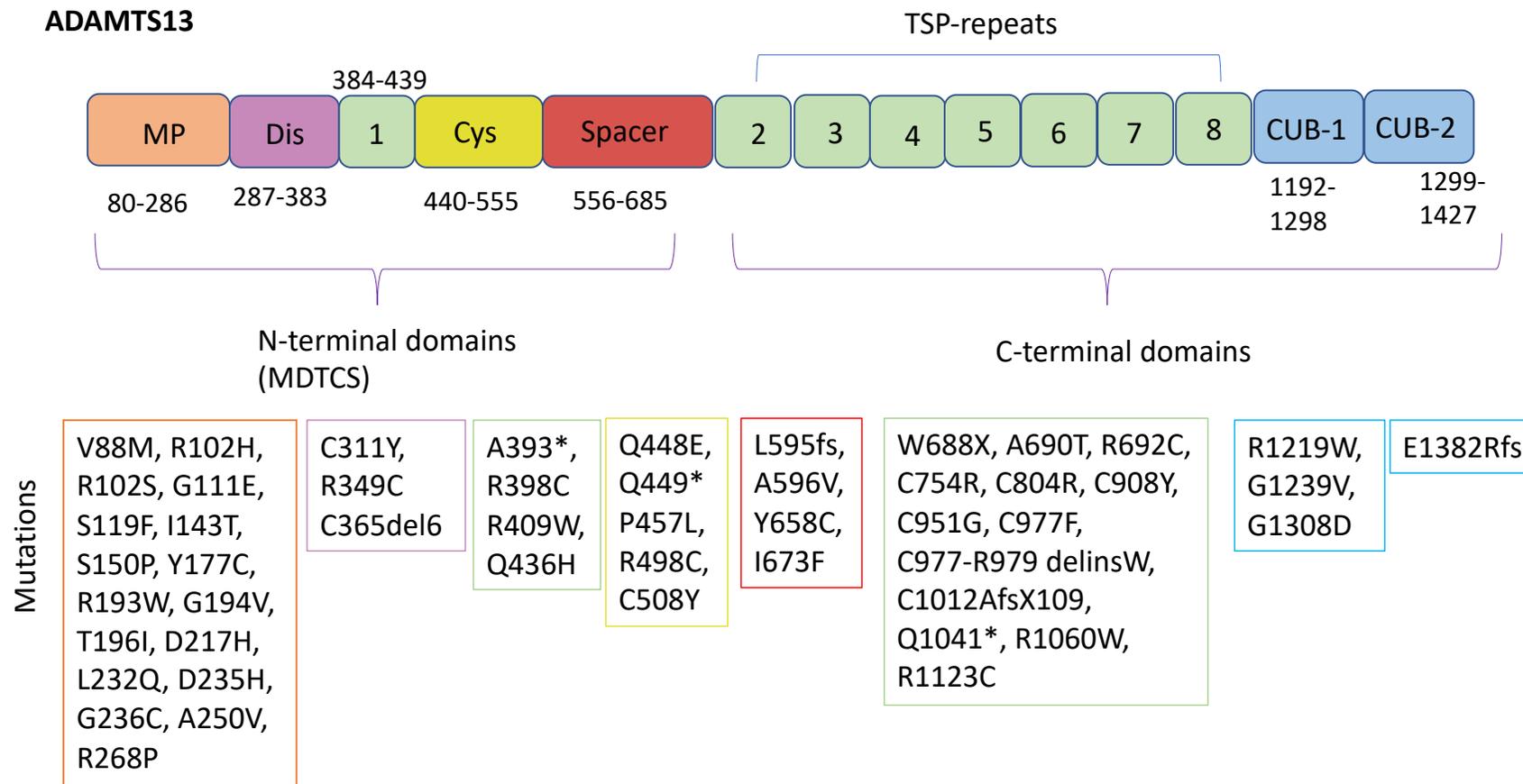


Figure 3.3 ADAMTS13 mutations by domain location. ADAMTS13 domain structure as seen in Figure 1.2 with mutations from Table 3.1 grouped by domain region. Mutations were included where evidence of likely pathogenicity (with TTP phenotype) was available from literature.

3.3 Discussion

Combining associations identified in multiple review papers as well as both the variant data and TTP mutation table enabled progress in elucidating the link between mutations and clinical features of TTP. This chapter highlighted some residues critical for ADAMTS13 and VWF interaction, as loss of these residues and subsequent interactions can result in disease. This analysis utilised the Ensembl database as well as literature to obtain patient mutation information, which was similar to previous papers (38, 189, 190). The review by Rizzo et al (2012) also reported a similar percentage of missense mutations (60%) with the remaining 40% composed of nonsense, frameshift and splicing mutations which were also common types reported in Figure 3.1B (191). This understanding of mutation types does not reveal anything further on ADAMTS13 in disease with many still having unclear pathogenicity. When collating TTP patient information, there was often insufficient records to provide a complete overview of a patient's clinical picture with supporting *in vitro* experiments. There is an ongoing study to record all patients diagnosed with cTTP into the UK TTP registry (167). The consistency in recorded information in the registry when the study is complete, should benefit further investigation into clinical associations of mutations of the ADAMTS13 gene with specific aims to advance the understanding of TTP mutation and disease associations. The analysis in this chapter was unfortunately limited in the number of mutations without this UK TTP registry information, instead aiming to highlight areas in which more research was needed. Interrogating the available data has highlighted that it is hard to ascertain disease severity associations due to bias in initial reporting of patients in literature (before the registry was introduced with consistent reporting) where records were incomplete or focused on patients with something notable in their disease course.

The onset of disease reported in Table 3.1 was simplified into neonatal/childhood or adult onset to align with the 2 presentation peaks suggested in Alwan et al 2019. It was reported 69% of adult presentations were associated with pregnancy which was also a noticeable theme in this analysis (reported across ADAMTS13 domains) (Table 3.1). The review by Alwan also gave the first indication of a link between mutation location with clinical disease course, associating prespacer mutations and earlier onset disease. No other reviews have reported this association thus far and

therefore analysis should be undertaken on a larger cohort to confirm this link (38). It would be of interest to include information from multiple TTP registries to get the largest cohort of mutations to enable confident associations between mutation groups, following introduction of TTP registries in multiple countries (192). As Van Dorland et al (2018) highlighted, currently forming a diagnosis based only on mutation information does not give any indication of disease progression (193). Once treatment for an acute episode is finished, the strategy is often ‘watch and wait’ as prophylactic treatment does not give sufficient benefit to all patients (64). Assessing patient’s relapse likelihood based on mutation and structural information would help with personalisation of treatment courses to benefit more patients.

In vitro experiments have been critical in improving understanding of abnormal ADAMTS13 activity and secretion to elucidate the phenotype of heterozygous mutations. Previously only homozygous mutations could be linked to specific TTP manifestations or disease courses, and multiple review papers subsequently carried out experiments to assess defects in secretion and activity. The differentiation between secretion and activity defects would not be detected by the routine ADAMTS13 activity assays which do not detect specifically if the protein is in fact not present or not functional. There is no evidence thus far that suggests if either secretion or activity leads to a more severe disease course, but this remains an area for potential associations to be investigated with a large patient cohort. Analysis of the TTP mutations in Table 3.1 highlights the influence of structural perturbation of ADAMTS13 on the dysfunctional activity or secretion. Mutations at the active sites and exosites affect ADAMTS13 activity likely through issues with VWF recognition and cleavage, whilst destabilising mutations (such as truncations or interruption of essential domain interactions) can lead to reduced or abolished secretion, indicating a link may exist between precise mutation location and protein function.

Recent studies investigating the structural effects on ADAMTS13 caused by certain mutations have highlighted the importance of structure prediction for understanding disease severity. A summary of research into structural modelling for disease prediction is given in Wang et al 2016, where structure-based assessments on protein-ligand interactions are utilised across human disease to assess potential pathogenicity of mutations (194). Numerous reported ADAMTS13 mutations have

no clear evidence to aid in understanding on the structural effect on the protein, evidenced by the high number of ‘unclassified’ REVEL scores (Figure 3.2A). There are still gaps in the knowledge of ADAMTS13 latency and activation as well as the specific interactions that occur with VWF during cleavage. Pathogenicity prediction software rely on detailed experimental structures of proteins. Achieving experimental active structures of both ADAMTS13 and VWF would be highly beneficial, revealing information essential amino acids for normal VWF binding and cleavage by ADAMTS13. Consequently, modelling of mutations on these experimental structures by prediction software may improve predictions of pathogenic mutations leading to TTP. Interrogation into critical residues for normal functioning and stability of ADAMTS13 could also help drive TTP diagnosis and treatment basing guidelines on sound understanding of normal haemostasis and the pathogenic changes that occur in disease.

Whilst ADAMTS13 mutation information can’t currently provide information on clinical manifestations of TTP; research investigating ADAMTS13 structure, activation and stability are helping to bridge the gap in genotype-phenotype understanding. Although mortality rates are much improved in the recent years, many patients are still experiencing severe relapses and courses of disease. Making further progress in efficient diagnosis and treatment of TTP are essential to improving patients’ quality of life. Advancing understanding of the genetic perturbation of ADAMTS13 in TTP, and subsequent disease pathophysiology and progression, should improve both diagnosis and disease progression information for patients whilst also providing foundations for personalised treatment pathways in the future. A key link to bridging mutation information and improvements in patient diagnosis and treatment is understanding the structures of ADAMTS13 and VWF in both normal and disease states, which will be the focus of the subsequent research carried out.

4. Prediction of VWF-ADAMTS13 interactions to inform construct design for structural studies.

4.1 Introduction

X-ray crystallography has been utilised since the 1950s for resolving of chemical and biological structures, with a wide range of utilities including drug design, interrogating complex interaction pathways, and understanding ligand-receptor thermodynamics, to list a few (195, 196). To generate a crystal structure suitable for 3D structure determination of a protein, planning must start many steps previous to the X-ray diffraction experiments. The process begins with generating a protein construct suitable for expression, purification, complexing, and finally crystal formation, which can give rise to many potential issues to foresee and subsequently address. Advances in structural biology technologies have improved this process including novel model generation techniques as well as data processing pipelines (138, 140, 197). Utilising previous crystal structures of the target protein or finding homologs in the PDB can make these processes considerably easier, however recent developments in techniques such as AlphaFold provide a valuable tool to complement the development of constructs for crystallisation where previous limited information was present in the PDB (141).

Multiple VWF A2 domain fragments have been developed, of varying length, for use in assays (97, 198). VWF-96 encompasses the residues G1573-R1668, and VWF-73 is composed of residues D1596-R1668 both fragments target exosites at the known binding sites across MDTCS domains (previously described in Chapter 1.5) (97, 109). Thus far these fragments have not been crystallised, due to the nature of the open unravelled formation required to bind ADAMTS13 providing a state that's unfavourable for crystallisation alone. To be physiologically relevant, the constructs must be able to mimic the state of both proteins once activated following shear stress (9). Subsequently, when designing the protein fragments, it is important to ensure the construct will be stable enough for high level expression and purification and is conducive to stable complex formation. Whilst assays often only require a single binding, cleavage and release event, the complex formation for crystallisation must be stable enough to form crystalline structures suitable for X-ray diffraction. A

frequent method trialled to improve solubility, stability, and crystallisation of peptides, such as VWF, is through the addition of a fusion tag (199, 200).

Tag development for crystallisation is still an evolving topic and design needs to consider the specific space around, and interactions present within, a protein complex to ensure the resultant structure is still biologically relevant. As well as the tag structure, the removal of any peptide tags should also be considered as the presence of the protein tag may hinder crystallisation. To combat this problem crystallisation tags have been developed to promote crystal formation either through increasing the surface available for crystal contacts or covalently linking proteins to promote their interaction both having produced successful results. (199, 201). The small structure of the VWF peptide means that a fusion tag is essential for successful expression and purification. Considerations must be made first to any effect on tag size and placement in relation to the VWF-MDTCS complex. Furthermore, the effect of the tag on crystal contact formation should be considered, exploring a tag to promote crystallisation will be essential to the construct design.

Crystal structures of other ADAMTS proteins can provide useful insight into the metalloprotease activity of ADAMTS13. ADAMTS10 contains a propeptide domain which utilises a zinc trapping method (involving the interaction between cysteine and the zinc ion) between this propeptide and the metalloprotease domain to control latency. The presence of the cysteine switch was presented as a mechanism of inhibition for multiple metalloproteinases in 1990 and has since been investigated further in ADAMTS1 and 12 (202-204). Whilst all ADAMTS proteins are zinc metalloproteases, this cysteine switch latency control is not present in ADAMTS13. There is the least homology with other members of the ADAMTS family members to ADAMTS13, which cleaves off the short propeptide domain shortly after synthesis therefore has no control over secretion, folding or latency of the protein (26). The activation of ADAMTS13 is understood to be dependent on VWF binding (as previously mentioned in Chapter 1.6) but the full mechanism by which this occurs, and the precise residues involved are still unknown. Therefore, the development of new constructs for studying the interactions between ADAMTS13 and VWF in their active forms are required and this work will identify design

improvements including mutagenesis and tag design, for both successful interaction studies and crystallisation experiments.

The MDTCS domains have been previously crystallised in complex with a F(ab) fragment, this crystal structure provided revolutionary insight into the potential activation of ADAMTS13 and its interactions with VWF. The current structure available highlighted the global and local mechanisms of maintaining latency, however the full cycle of activation can't be fully understood until the structure of the active form of ADAMTS13 can be interrogated. The utility of a complex of VWF with the MDTCS domains is hoped to eliminate the use for the Fab required to stabilised MP domain in the previous MDTCS structure. Addressing all these issues will help generate a biologically relevant complex, suitable for crystallisation to elucidate the regulation of ADAMTS13 and VWF.

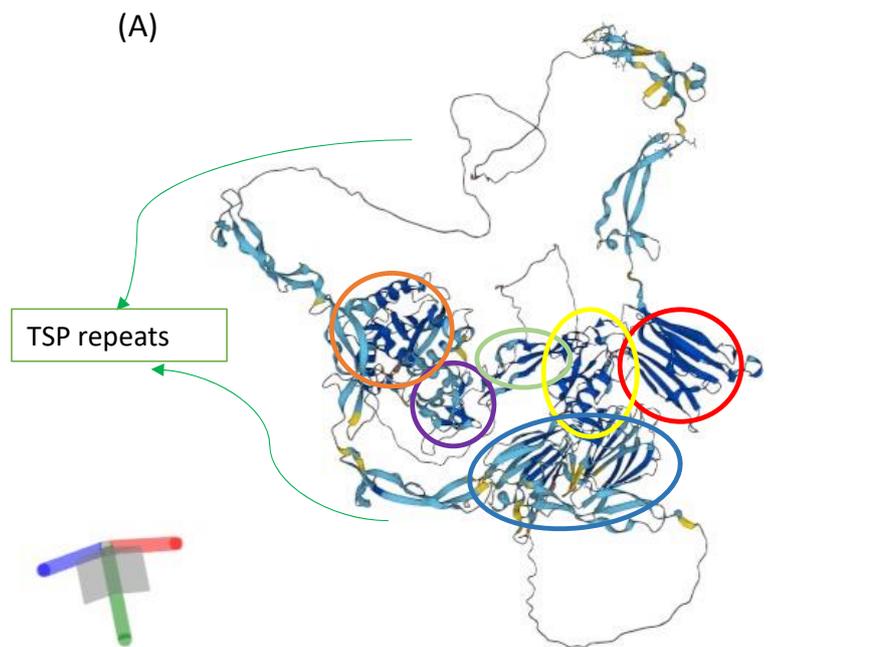
4.2 Results

4.2.1 Comparison of experimental and predicted ADAMTS13 structures.

An AlphaFold prediction of the full length ADAMTS13 protein is now available in the AlphaFold database and is shown in Figure 4.1A alongside the corresponding error score Figure 4.1B (ID: Q76LX8). A comparison between the AlphaFold model to the crystal structure of MDTCS (PDB: 6GIQ) was performed to evaluate the similarities and investigate the utility in the model for understanding the latent and active states of ADAMTS13 (Figure 4.2A). The MP and Dis domains appear to show large similarities with the position of alpha helix and beta sheets that make up these domains (Figure 4.2B). After this point differences in the structures can be seen. The hydrogen bonding present between the beta sheets in TSP-1 is maintained in the AlphaFold model however the angle at which this domain protrudes mean both TSP-1 and cysteine-rich domain occupies slightly different space in relation to the MP and Dis domains (Figure 4.2C). The cysteine-rich domain contains the same core beta sheet and alpha helix structures however their position relative to the neighbouring domains is different, as well as the AlphaFold prediction containing beta sheet loops that are missing for the MDTCS structure for amino acids (D454 to Y468 and I490 to K497). The spacer domain structure and hydrogen bonds are similar between structure and model, however the positioning of the spacer domain

in the AlphaFold model is much higher compared to the structure (Figure 4.2D). This could be due to the presence of the TSP repeats present in the AlphaFold model which are not present in the MDTCS structure. The CUB1-2 domains were recently resolved (PDB: 7B01) and AlphaFold model shows large similarities with the beta sheets that make up these domains (Figure 4.2E).

The TSP repeats 2-7 have not been previously resolved due to issues with their flexible and unstructured nature, as can be visualised in the AlphaFold model (Figure 4.1A). It's important to note the lack of confidence in AlphaFold LLG score for these domains as well as PAE score suggesting lower confidence in both the structure of domains and positioning relative to other domains (Figure 4.1B). The TSP repeats are predicted to form a large looping structure surrounding the MDTCS domains, forming 3 loose 'V-loop' structures. The positioning of the CUB domains in the AlphaFold model does not allow for an interaction with spacer domain that is supported in previous research (110, 124, 205). Instead, the CUB1-2 domains are in proximity to the TSP-1 domain but form no hydrogen bond interactions to other domains within the protein, indicating placement of these domains is unknown (Figure 4.1 and 4.2A and E). This highlights the need of improved modelling pipelines, to match the data quality achieved from generating crystal structures, to accurately investigate the activation states of a protein.



Model Confidence

- Very high (pLDDT > 90)
- High (90 > pLDDT > 70)
- Low (70 > pLDDT > 50)
- Very low (pLDDT < 50)

AlphaFold produces a per-residue model confidence score (pLDDT) between 0 and 100. Some regions below 50 pLDDT may be unstructured in isolation.

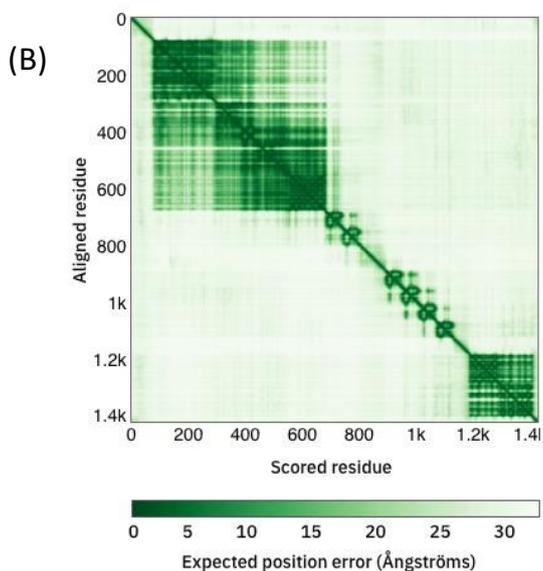


Figure 4.1 (A) The AlphaFold prediction for ADAMTS13 (ID: Q76LX8) with related model confidence estimates used for colouring. Domains are labelled using coloured circles (Orange: MP, Purple: Dis, Green: TSP-1, Yellow: Cys, Red: Spacer, Blue: CUB1-2, green arrows indicated the unstructured loops of the C-terminal TSP-repeats. (B) The expected error score given to each residue from the model seen in (A), more information on this score can be found at (145)

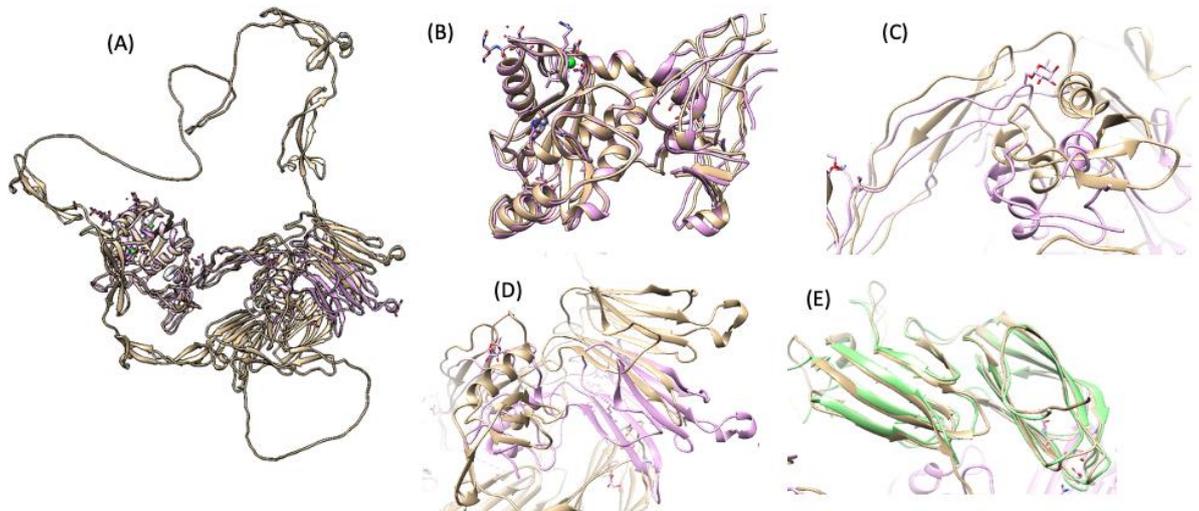


Figure 4.2 Aligned structures of ADAMTS13 experimental and predicted structures. (A) Aligned structures of ADAMTS13 (AF model in beige ID: Q76LX8) and MDTCS E225Q in pink (PDB: 6GIQ), indicating the agreement between MP, Dis domains but not the location of Cysteine rich and spacer domains. More details and a closer look at each domain in B-E, colours as in A with CUB1/2 (PDB: 7AOB) in green. (B) Metalloprotease and Dis domains showing good alignment (C) TSP-1 and Cys domains with differing relative locations between models (D) Spacer domain positioning differs between models (E) CUB domains align nicely when overlaid.

4.2.2 Guided docking of ADAMTS13 and VWF A2 domain

The AlphaFold prediction of the VWF A2 domain outputs a structure in the closed folded formation, not in a state amenable to ADAMTS13 binding (data not shown). To overcome the issues of the docking software being unable to predict the elongation of the A2 domain through replicating conditions of shear stress, a guided docking approach on the ROSIE server was utilised involving the MDTCS crystal structure (PDB: 6GIQ) and an elongated VWF structure based on the shape in Crawley et al 2011(9). The exosite interactions suggested from *in vitro* experimental analysis were utilised as a ‘guide’ for the molecular interactions between proteins; these residues are detailed in Table 1.1. The predicted structures and interactions present can be seen in Figure 4.3 A-B. The MP domain of the modelled MDTCS structure is in the closed state, indicated by the gatekeeper triad of residues interactions remaining intact, however the prediction does indicate the expected interaction of D252 with Y1605 and M1606 of VWF. This suggests some utility in

the predictions with the localisation of VWF to ADAMTS13 exosites from previous predictions.

The Dis domain exosite residue R349 interacts with D1614 as expected, however no interactions are seen with L350, instead the model indicated interactions around S346, C357 and S348 with I21 which lacks previous experimental evidence. Similarly, the interactions with Cys-rich domain residues focus around P475 and H476, instead of G471-A474, interacting with D58. The spacer domain interaction more closely aligns with expected interactions, involving R660, Y661 and Y665 with VWF residues Q72, R73, L71, D68 and L69. The predictions for interactions can therefore provide models of localisation but exact interactions may not be accurate. This suggests utility of predictions informing construct design but not biologically relevant interaction information that can be obtained from complex crystal structures themselves.

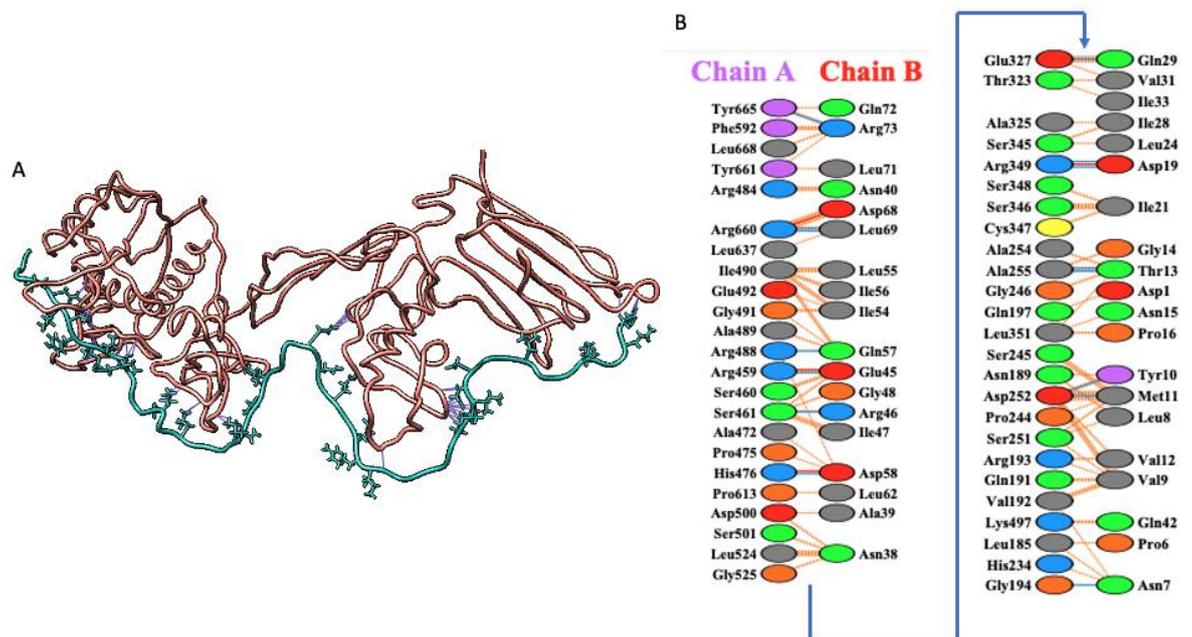


Figure 4.3 Guided docking between MDTCS E225Q (PDB: 6QIG) and VWF (Gly1573-Arg1668) represented as a schematic in (A) with MDTCS in red, and VWF in blue, interactions (below 3Ång) are shown as purple lines. (B) Ligplot schematic of interactions between MDTCS E225Q (chain A) and VWF (Chain B) Interactions are represented as lines with salt-bridges in red, hydrogen bonds in blue and non-bonded contacts as orange dashed lines

4.2.3 Design of VWF mutants to improve stable interaction with ADAMTS13.

To generate VWF constructs (based on the VWF-73 construct described previously) suitable for expression, purification and subsequently crystallisation, details from previous experiments were evaluated. The interaction information provided by Crawley (Imperial) indicated that the main issues with stable complex formation with ADAMTS13 were caused by VWF being in a state inaccessible to ADAMTS13 or the interaction not being strong enough to maintain a stable complex. So mutant VWF constructs were designed to overcome these issues encountered, with the 6xHis-SUMO-VWF construct utilised in Petri et al 2019, used as the WT VWF construct in the experiments for comparison (97).

To generate a stronger interaction between the VWF and ADAMTS13 proteins, the manipulation of the Zn^{2+} ion of ADAMTS13 was investigated as a potential mechanism. To understand the positioning of VWF within the metalloprotease domain, other ADAMTS proteins were considered alongside published literature detailing essential interactions for binding and cleavage of VWF by ADAMTS13. The AlphaFold database was used to locate AlphaFold models of both ADAMTS4 and 10, and models were aligned focusing on the metalloprotease domain (ID: Q5VTW1 and M0QY12). The distance between the zinc ion and cysteine residues for these proteins was between 3.7Å and 5.2Å for the cysteine rotamers predicted in AlphaFold models (Figure 4.4A). When aligning ADAMTS13 with these structures, it is indicated that the space occupied by the pro-domain of the ADAMTS family proteins may in fact be replaced with VWF in ADAMTS13-VWF complex. The similarities between unravelled VWF and pro-domain are clear with the long linear single amino acid protein structure (Figure 4.4B). Current predictions do not place VWF within this cleft as seen in ADAMTS13 as well as ADAMTS4 and 10 (Figure 4.4C), however these structures give a good basis of a model for interaction and subsequent construct design by putting the scissile bond and active site in closer proximity. Mutation of the tyrosine and/or the methionine at the scissile bond should therefore place cysteine in the desired location to co-ordinate the Zn^{2+} ion and form the stable complex required (mutants named Y1605C and YM1605-6CC).

Another mutant construct was also created to address the issues with VWF not forming the open elongated state suitable for ADAMTS13 interaction, likely due to the lack of

shear stress experienced in the blood. The mutant named G1629E was proposed due to previous research indicating the success of the mutation for destabilising VWF and improved ADAMTS13 binding (206). The G1629E mutant is located in the H5 loop, and results in a mechanical perturbation of the whole A2 domain resulting in an open linear structure (Figure 4.5). This can't be predicted when visually comparing the WT VWF A2 domain (PDB: 3GXB) with the modelled G1629E mutant, but instead relies on simulations to replicate the environment of shear stress as seen in the blood and these experiments were carried out previously (206, 207). This research did not directly conclude which loss of interactions and structural rearrangements lead to this destabilisation, but it could be predicted the large side chain of Glutamic acid compared to Glycine, as well as the charge, may mean the local area surrounding the amino acid is altered, thus affecting the H5 loop. This may explain the preference for G1629E VWF peptide towards the open unravelled confirmation (Figure 4.5).

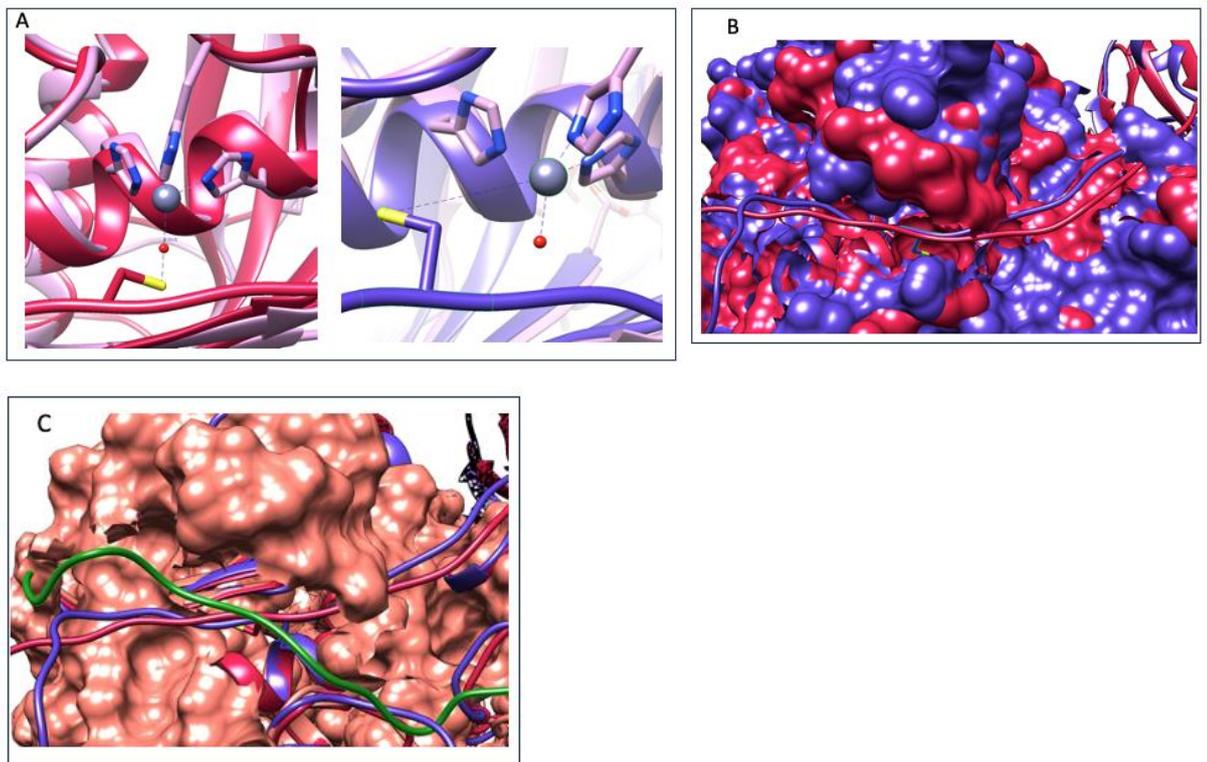


Figure 4.4 Modelling and docking of VWF mutant constructs. (A) AlphaFold structure of ADAMTS 4 and 10 in (dark pink and Purple respectively) with MDTCS E225Q (PDB: 6GIQ) (in light pink) illustrating the proximity of the cysteine residues in the prodomain of ADAMTS4 and 10, to the Zn²⁺ ion of ADAMTS13 when structures are overlaid. The proximity of cysteine and Zn²⁺ in AD4 and AD10 is sufficient for interaction with the prodomain. (B) ADAMTS4 and 10 (Colours as in (A)) are here represented with the MP

domain surface shown to create a long cleft for which the prodomain to snake through (depicted as the wire structure). (C) MP domain surface of ADAMTS13 (6GIQ) in peach shows a similar cleft present to that of AD4 and 10, and when overlaid these AD4 and 10 prodomains lie within the cleft (colours as in A) whereas the current prediction of VWF is not accommodated in this cleft (green wire), highlighting issues with current structure prediction.

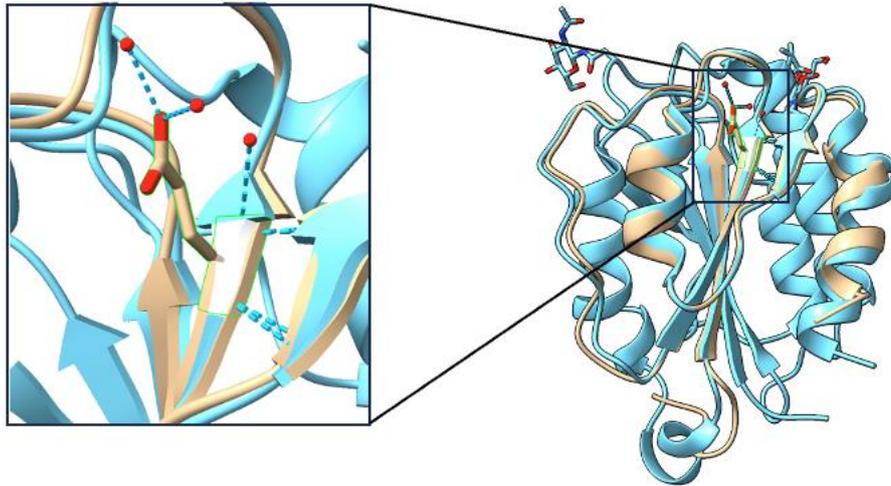


Figure 4.5 Cartoon representation of the VWF A2 domain comparing WT and G1629E mutant. The VWF A2 structure is shown with the WT in blue and the G1629E mutant in beige. The zoomed profile illustrated the lack of carbon side chain of glycine (white section of ribbon) compared to the bulky side chain of glutamic acid which takes up a much larger amount of space and also interacts with a further water molecule. This large size may lead to the destabilisation noted in the H5 loop which subsequently perturbs the entire domain.

4.2.4 Generation of 6xHis-SUMO-VWF73 mutants for complexing with ADAMTS13

For both mutants, primers were designed for use in PCR mutagenesis experiments, primer sequences and construct sequences can be found in Chapter 2.2. The double mutant containing both mutations was also created by utilising both sets of primers leading to production of 6xHis-SUMO-VWF73 constructs named WT, Y1605, G1629E, Y1605C+G1629E and YM1605-6CC. PCR experiments were run and after transformation of resultant mixture in NovaBlue *E.coli*, sequencing was utilising T7F and T7R primers were used to confirm successful mutagenesis of VWF construct sequences (Supplementary Table 1).

4.2.5 2GKG construct design.

In order to increase the likelihood of successful crystallisation, alternative VWF constructs were generated utilising the 2GKG tag which previously aided in successful crystallisation of challenging proteins (208). The location of the tag and VWF peptide in relation to each other and ADAMTS13 were considered to generate meaningful interactions whilst also improving expression and stability of the constructs for complex formation. The VWF peptide was also altered in an attempt to improve both interaction and subsequent structural studies. The VWF peptide was shortened so one construct contains 2GKG fused to cysteine-rich and spacer domain exosites of VWF, and other constructs contain 2GKG fused to Metalloprotease and disintegrin domain VWF exosites, with a final construct consisting of 2GKG inserted between the two exosites mentioned previously (Schematic found in Figure 2.1 and sequences in Supplementary Table 1).

To generate these constructs, AlphaFold predictions were utilised to provide predictions about the interaction between the VWF peptide and MDTCS domains to which they interact with, these complexes can be seen in Figure 4.6A-C. Two constructs were generated utilising the 2GKG tag and MP and Dis domain exosites, containing a shorter or longer linker between (constructs named 2GKG-MD-S or 2GKG-MD-L respectively). It was unclear from the predictions which linker length would best facilitate binding, so both were utilised to increase crystallisation success. Limited evidence thus far for any involvement with TSP1 domain for ADAMTS13 binding meant the 2GKG tag could be placed between known exosites without losing any critical interactions to VWF (Figure 4.6A). The locations of 2GKG-VWF fragments in relation to ADAMTS13 matched that witnessed with the MDTCS-VWF model in Figure 4.3, the direct interactions were not studied closely around the MP domain which is still modelled in the inactive form so instead this general localisation of VWF exosites to the ADAMTS13 MP and Dis domains provided sufficient confidence for the approach. The 2GKG-CysSpacer construct with the 2GKG tag and Cys-spacer exosites linked was modelled with the Cysteine rich and Spacer domains of ADAMTS13, also showing good localisation between proteins and similarities to the previous model (Figure 4.6C and 4.3)

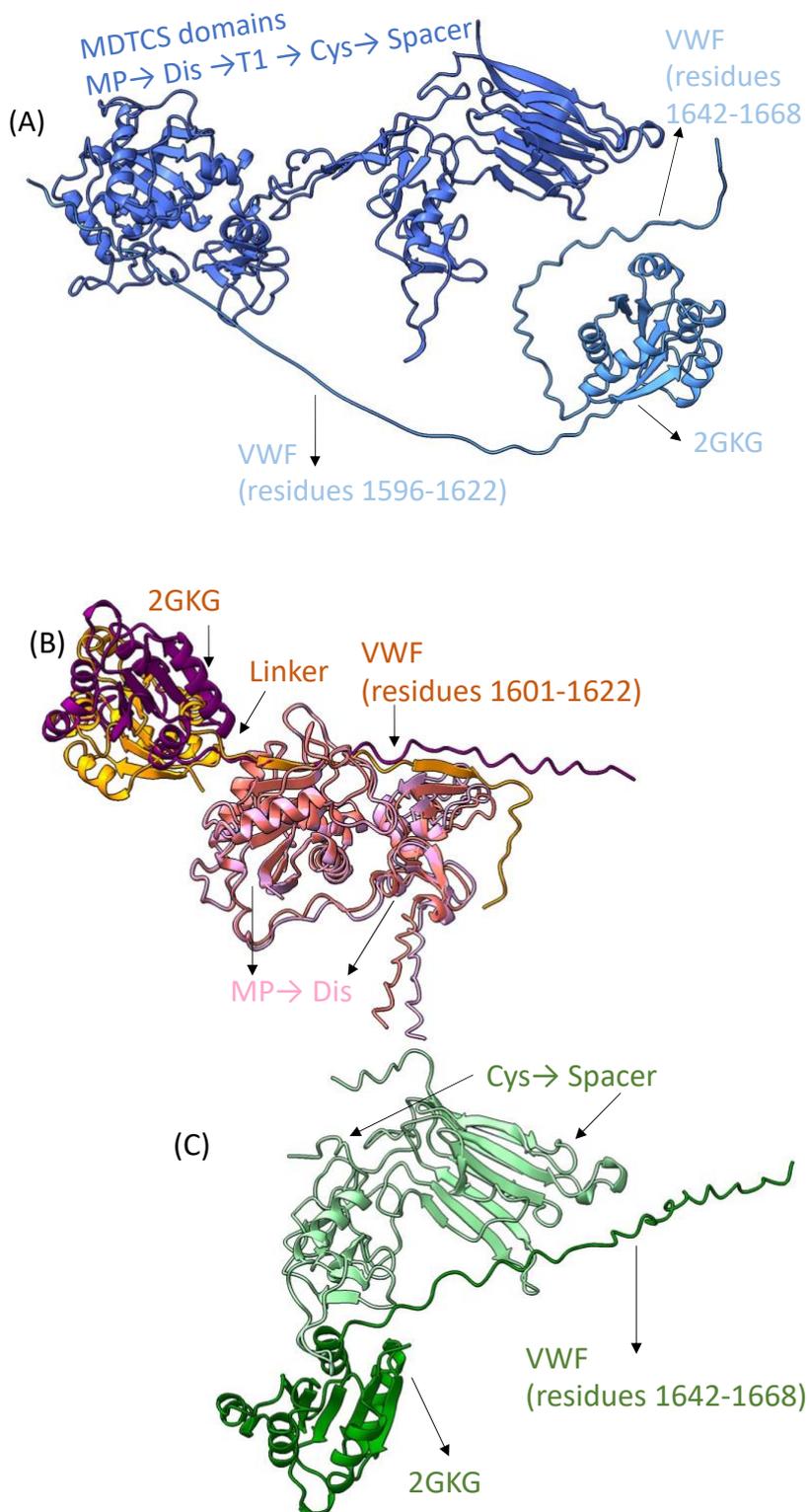


Figure 4.6 AlphaFold modelled structures of VWF-2GKG constructs docked with ADAMTS13 domains. (A) 2GKG-VWF-Internal in cornflour and MDTCS in blue (B) 2GKG-VWD-MDs in orange and 2GKG-VWD-MDI in purple with MD domains (C) 2GKG-VWF-CysSpac in dark green with CS domains of ADAMTS13 in light green.

4.2.6 Generation of 2GKG-VWF constructs for complexing with ADAMTS13

The 2GKG-VWF constructs were prepared by gene synthesis which were PCR amplified for use in restriction cloning into the pet-25b vector. Cloning was performed in collaboration with a lab member utilising NheI/XhoI and the correct insertion of the fragment was confirmed by sequencing (Supplementary Table 1).

4.2.7 ADAMTS13 construct design for crystallisation studies.

For ADAMTS13 constructs, previous research was studied and the MDTCS E225Q construct was gifted by J. Crawley (Imperial) as well as the WT ADAMTS13 full length (FL) construct, sequences can be found in Supplementary table 1. As well as these constructs, another was designed in collaboration with a lab member. The DTCS construct is designed based on the minimal functional ADAMTS13 protein, suggested in Zhu et al 2019 as the *Columba Livia* (pigeon) protein lacking TSP3-6 (12). This smaller construct whilst allowing interaction between the Spacer-CUB domains, contains a shorter sequence of TSP repeats which should aid in the issues of flexibility in this region which would likely crystallisation. The DTCS-CUB construct contains DTCS domains followed by T2,7-8 and CUB1/2 domains, sequence in Supplementary table 1. An AlphaFold model for this construct was also generated (Figure 4.7A), the TSP-1, cysteine-rich and spacer domains show good agreement with the MDTCS structure, with just the Dis domain slightly misaligned in comparison (Figure 4.7B). Interestingly, compared to the ADAMTS13 AF prediction in the database, the TSP repeats (2-8 for FL and 2,7,8 for DTCS) and CUB1/2 lie in a different location relative to the MDTCS domains (Figure 4.7C).

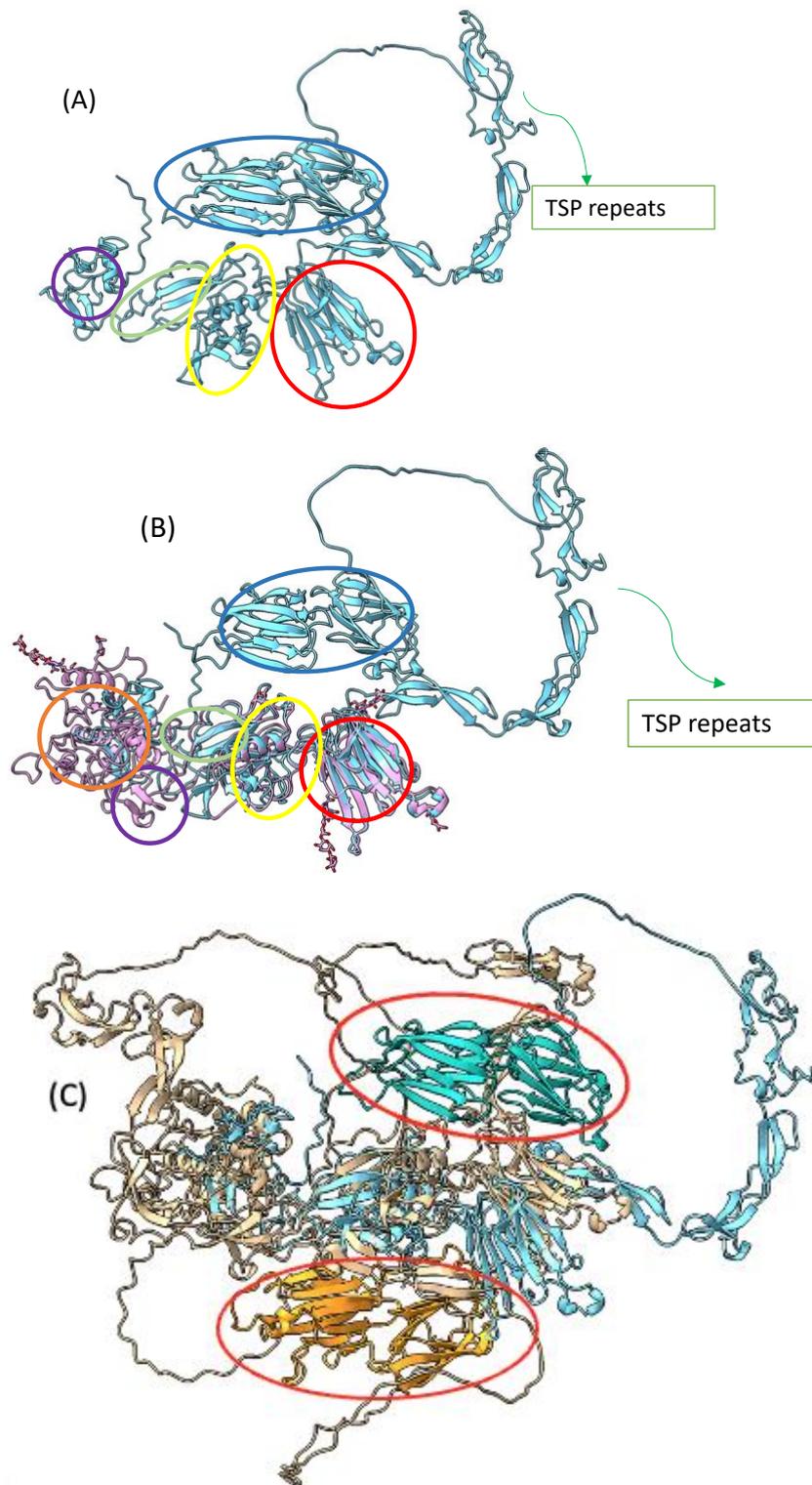


Figure 4.7 AlphaFold predictions of DTCS and comparison to MDTCS experimental structure and predicted ADAMTS13 FL predicted protein. (A) The AlphaFold predicted DTCS structure shown in blue, with the TSP repeats allowing the folding back of the CUB domains to within proximity of the Spacer domain. Domains are labelled using coloured circles (Orange: MP, Purple: Dis, Green: TSP-1, Yellow: Cys, Red: Spacer, Blue: CUB1-2, green arrows indicated the unstructured loops of the C-terminal TSP-repeats (B) The

MDTCS E225Q structure (pink) is aligned with DTCS (blue) to illustrate the similarities with the TCS domains when structures are overlaid. Domain colours as in (A) (C) When the ADAMTS13 FL structure (beige) is overlaid with DTCS (blue) the differences in precise domain location is clear, as well as the location of the CUB domains (illustrated in brighter colouring of each chain as well as a red circle) with neither ADAMTS13 FL or DTCS AlphaFold predictions localising the CUB domains to the spacer domains, likely due to the flexible and unpredictable nature of the TSP repeats for AlphaFold

4.3 Discussion

In order to successfully investigate the interaction and regulation of VWF by ADAMTS13, new VWF constructs were required to generate a stable complex with ADAMTS13 suitable for crystallisation. Constructs were generated utilising information from molecular docking and AlphaFold structures on other ADAMTS family proteins as well as VWF.

The AlphaFold structure predictions provide a basis for developing constructs based on predicted interaction information and molecular structure. Interactions were present in the modelled VWF-MDTCS complex in both areas of expected contact within exosites and also between residues not previously explored in detail as sites for VWF binding. The spacer domain interactions with VWF match evidence from the literature involving Arg660, Tyr661, and Tyr665 as a cluster of residues, which are also site of autoantibody binding due to the surface facing exosite location (209).

The Disintegrin and Cysteine-rich domains however did not completely agree with previous published structures, indicating the alignment of VWF with these residues may be slightly off. These differences from expected interactions may be explained by the closed state of ADAMTS13 modelled, the MP domain is not in a state to accommodate the VWF even with interactions specified in the software; indicated by the gatekeeper triad remaining intact. It's possible a shift in the position of the VWF residues may occur when the peptide is accommodated by ADAMTS13, as well as the structural rearrangement needed for ADAMTS13 to facilitate the interaction. Evidence for this change was highlighted previously with comparison to ADAMTS1,4 and 5 active sites where the structural rearrangement required can be

visualised (97). There were also further interactions between ADAMTS13 and VWF evident in the model (Figure 4.3) which suggest there may be other interactions present that have not yet been explored in previous research. However, no conclusions can be drawn from this prediction without further evidence from either *in vitro* work or crystallisation experiments, an experimental model is required to support or refute features of the prediction. This again highlights the importance of generating constructs capable of interacting to form a stable complex so the structure of open active states of these proteins can be resolved.

The AlphaFold output models generated indicate insufficient information of protein activity states and should therefore be used only as models in this scenario. The modelling of the open and closed states of proteins is not yet developed enough to provide an indication of interaction especially in the MP domain of ADAMTS13. The location of domains in proximity to another is also another area that requires improvement, indicated by low PAE scores for ADAMTS13 (Figure 4.1). All AlphaFold models generated in this chapter provide some regions of difference in their structure prediction, especially around TSP repeats and CUB domain placement. This could be merely due to sequence differences, such as the number of TSP repeats in DTCS compared to FL ADAMTS13, but conclusions cannot be formed with little support for one prediction over another. Furthermore, the modelling software is unable predict the active forms of ADAMTS13 and VWF, without input such as from simulation studies to create an interactive model. Whilst predictions can provide utility in certain scenarios to guide experiments, such as in this case with design of constructs, the main method for structure resolution remains crystallography and the subsequent interrogation of experimental structures.

Nonetheless, the information available for the family of ADAMTS proteins in the AlphaFold database was exceedingly useful for construct design. These predicted structures provided structural information of the propeptide domain, important for activation of some ADAMTS proteins, that matches previous research and thus was beneficial for generation of the VWF Y1605C mutant. This Y1605C mutant hopes to provide a more stable interaction between ADAMTS13 and VWF, and as the cleavage of the scissile bond in this context is unfavourable no issues are expected from the mutation at this site. The use of these ADAMTS family proteins was also

explored by Petri et al (2019), previously around the MP domain latency and VWF accommodation, providing support for the utility in guiding construct design for this VWF mutant (97). The approach of co-ordinating the metal ion has not been utilised previously for crystallisation, but stable complex formation is often required for inhibition of metalloproteases such as those summarised by Jacobsen et al (2010), so has been explored to some degree previously (210). The G1629E mutant was developed to enable VWF to take the open elongated conformation suitable for ADAMTS13 binding. As no crystal structures are available for this form, previous research using energy calculations, and the hypothesised mechanism of destabilisation were utilised for construct design. It's hoped the short VWF peptide will follow the same mechanisms of the VWF A2 domain, and as the modelled prediction suggests the same tertiary structure for VWF some support is provided. Mutations to destabilise proteins for complexing is not a developed area of research, likely due to the unique nature of interactions and activation VWF and ADAMTS13 have. However, it is clear that proteins are often destabilised through disease causing missense mutations such as G1629E found in Von Willebrand factor disease type 2A, as well as many other human diseases (211, 212). A final 6xHis-SUMO-VWF73 mutant will encompass both these mutations in an attempt to address the problems of stable complex formation between the proteins.

To provide an alternative strategy for complex crystallisation, the 2GKG-VWF group of constructs were also developed. Fusion tags each have their own benefits and draw backs, with the SUMO-tag aiding in expression and previous research with VWF but with limited data available on crystallisation success. For this reason, the 2GKG tag was utilised with benefits of expression and crystallisation but does lack any previous research with VWF or ADAMTS13. The 2GKG-VWF constructs are designed to aid in crystallisation through maintaining exosites essential for interaction with ADAMTS13 whilst removing flexible regions that may inhibit the crystallisation process. The design of these constructs utilising separate complexes to visualise specific exosite interactions occurring between ADAMTS13 and VWF may not generate the entire picture of activation cycle but will help provide a basis for understanding. Thus far utilising *in vitro* experimental information was useful in guiding docking of all VWF and ADAMTS13 constructs utilising residues located in exosites known to be essential for interaction.

The constructs have been successfully cloned and amplified and are suitable for continuation with research into their expression in the respective systems for production of ADAMTS13 and VWF proteins. The breadth of constructs should enable progress in exploring the best method for interrogation of interactions between ADAMTS13 and VWF. Furthermore, the utility of AlphaFold has been explored with an emphasis on the points at which during structure solution it may best be utilised, such as in construct design, but the limitations for biologically relevant information highlight the need for experimental structure solution. This is especially relevant in complex cases such as the ADAMTS13-VWF complex where activation states of the proteins are crucial to understanding the proteins latency cycles and the effect this plays on their interaction and complex formation.

5. Exploring complex formation and the interactions between ADAMTS13 MDTCS and VWF A2 domains.

5.1 Introduction

Thus far ADAMTS13 structures deposited in the PDB have been expressed in two expression systems: mammalian, using CHO cell lines, and insect using drosophila melanogaster cell lines (PDB codes: 3GHM, 3GHN, 3VN4, 6QIQ, 7B01).

Expression in mammalian cells is preferred to ensure the complex folding of the protein is correct, facilitated by the extensive glycosylation network which these cells are capable of replicating. One downfall of using this system however is the incredibly low protein yield. Whilst $\mu\text{g/ml}$ amount may be suitable for some activity assays and other experiments, protein crystallisation experiments require a much higher concentration of protein ($\sim 10\text{mg/ml}$ is often initially trialled) (213). Attempts have been made to improve mammalian expression as seen in Kagro et al 2022 and Jankowzaka et al 2022 however these expression levels are still suboptimal for crystallisation and come with high costs for producing the amount of culture media required to scale up protein concentration(214, 215).

For the DTCS fragment (PDB codes: 3GHM, 3GHN, 3VN4), purification in CHO cells was successful in generating protein concentrations suitable for crystallisation (109, 216). The DTCS fragment only contains 4 glycosylation sites whereas the metalloprotease domain of ADAMTS13 contains a further 2 glycosylation sites meaning expression and subsequent crystallisation are more challenging (217). Glycosylation can hinder crystallisation due to the bulky, heterogenous, mobile sugars present; investigation into optimal conditions can be required(218). Previous reports have highlighted the challenges with expression of metalloproteases, including requirements for specific buffers and conditions, such as calcium to maintain activity of some ADAMTS proteins including ADAMTS13 (219). The most recently published ADAMTS13 structures (namely PDB: 6QIG) utilised the insect cell system for protein overexpression, with benefits of reduced costs and time required from induction to harvesting of protein. The glycosylation present on MDTCS (PDB: 6QIG) was consistent with previously reported literature, giving

support for this expression method and the successful production of the MDTCS construct(97, 217).

The MDTCS structure (PDB: 6QIG) provided a ground-breaking insight into the previously unclear mechanisms of activation. The elucidation of the local mechanism (gatekeeper triad) and global mechanism (Spacer-CUB interaction) of inhibition, clearly highlights the conformational sensitivity of MDTCS activation (97, 110) Differences were identified between the structure of DTCS (PDB: 3GHM and 3GHN) and MDTCS including the twisted and declined positioning of the Dis domain relative to the Cys/Spacer domains (97). To gain more detailed information on the gatekeeper triad and calcium loops of the metalloprotease domain and how these are manipulated in the activation of ADAMTS13, more research is needed.

Research utilising antibodies or VWF peptides to disrupt the Spacer-CUB interaction have provided support for the ‘folded’ or ‘closed’ formation mediated by this interaction and subsequent hiding of the spacer exosite; further supported by SAXS data(124, 125). Disruption of this interaction leads to the ‘open’ formation providing a more linear global structure, as well as local changes within the MP domain (more details given in Chapter 1.6) (220). Further interaction points have been suggested by De Young et al 2022 involving the TSP1-5 and MP, TSP1-7 and disintegrin, TSP1-8 and Cysteine-rich, as well as CUB-2 and disintegrin domain indicated by AlphaFold2 prediction. However these results refute previously identified sites of interaction from docking and have no experimental evidence as support so should be interpreted with caution(221).

To support or refute these theories, a complex crystal structure of ADAMTS13 and VWF is required. Construct design was covered in Chapter 4; to confirm if these constructs are suitable for ADAMTS13-VWF interaction and crystallisation studies, further experiments were carried out to characterise these constructs and begin crystallisation attempts. Experiments utilising isothermal calorimetry and surface plasmon resonance analysis will provide preliminary data on the interaction strength between VWF and ADAMTS13 complexes supported by size exclusion chromatography. Initially size exclusion analysis can confirm if complex formation is taking place. To support this, information from ITC provides details on binding

modes and ratios through slight changes in temperature whereas SPR gives binding and dissociation kinetics values through very slight changes in mass. Information about this interaction will guide subsequent complex formation and improve success in crystallisation trials.

5.2 Results

5.2.1 Protein expression and purification

5.2.1.1 VWF constructs

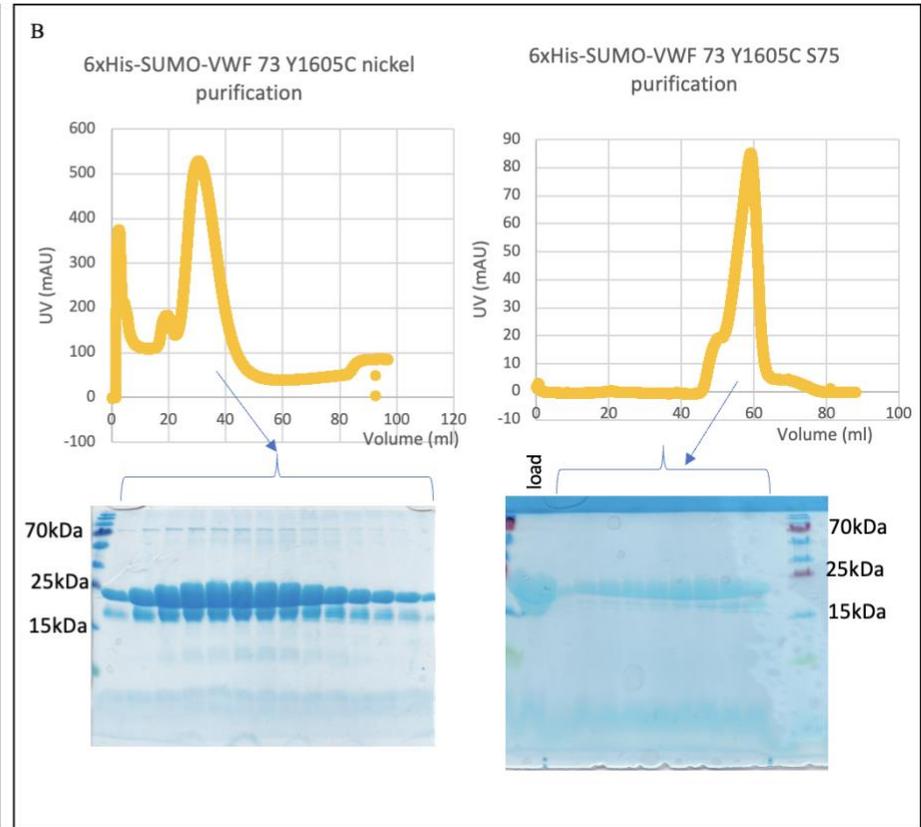
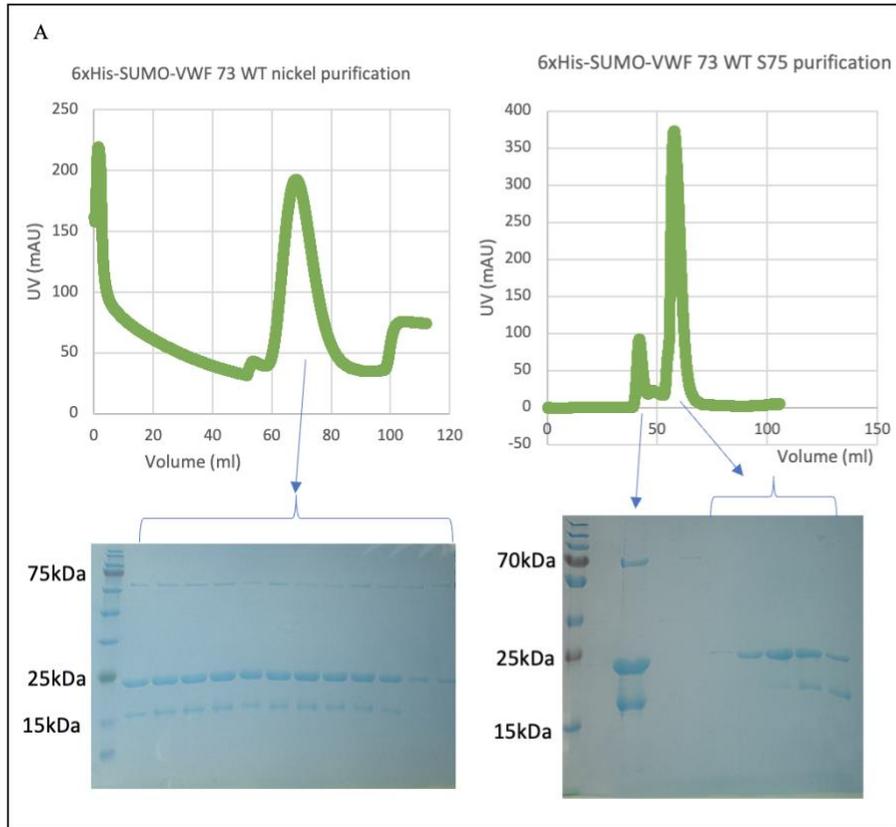
6xHis-SUMO-VWF was designed with a 6x-His tag to enable purification using a nickel affinity column as a first step after protein overexpression. 6xHis-SUMO-VWF constructs (WT, Y1605C and G1269E) eluted from the Ni column at a similar concentration of imidazole (100mM) elution profiles and SDS-page gels of the fractions can be seen in Figure 5.1A-C. Initially a Capto-Q column was utilised next for an intermediate purification step as utilised Petri et al 2019 (97), however there seemed to be limited binding to this column and protein loss outweighed any benefits to purity (Supplementary Figure 1). For this reason, the second and final purification step utilised was size exclusion on a Superdex S75 column (Figure 5.1A-D). Further improvements were made to VWF purity by running a slower flow rate (0.5 rather than 0.8ml/ml) to improve the separation achieved over the length of the column (Supplementary Figure 2).

The overlaid elution profiles from gel filtration can be seen in Figure 5.1D, with all 6xHis-SUMO-VWF constructs eluting at similar volume- as expected due to the same MW. The slight difference in elution volume seen with G1269E could be due to the instability provided by the mutation to aid complexing with MDTCS E225Q. Although all constructs are eluting at the same volume, a MW of 22kDa would expect to elute at ~80ml not the ~60ml observed according to the Cytiva manual which required further investigation. The SDS-page gels for each construct can be seen in Figure 5.1A-C. Although the elution peak suggests a pure protein product, there is evidence of some degradation with lower molecular weight species occurring as visualised on the gel within this elution peak. This protein was flash frozen and taken forward to further experiments even with this unexpected degradation, to

confirm the identity with western blot and cleavage analysis as well investigating any effect on complexing experiments with MDTCS E225Q.

The construct Y1605C+G1629E produced a protein that was unstructured and did not express to a high level so pure protein could not be obtained following purification (Supplementary figure 3) and construct YM1605-6CC did not express (data not shown) so both constructs were not carried forward for further experiments.

VWF-2GKG constructs were subsequently purified following the finalised method for 6xHis-SUMO-VWF protocol. All constructs eluted from the Ni column at a similar concentration of imidazole (85mM); elution profiles and SDS-page gels of the fractions can be seen in Figure 5.2A-D. The final purification step utilised was size exclusion on a Superdex S75 column, elution profiles can be seen in Figure 5.2, with an overlay of the elution profiles included (Figure 5.2E). Whilst only one peak is observed on the trace, on the SDS-PAGE some higher weight species of VWF can still be seen (Figure 5.2). This main peak was still collected and taken forward for analysis by western blot to confirm all protein collected is pure VWF-2GKG constructs.



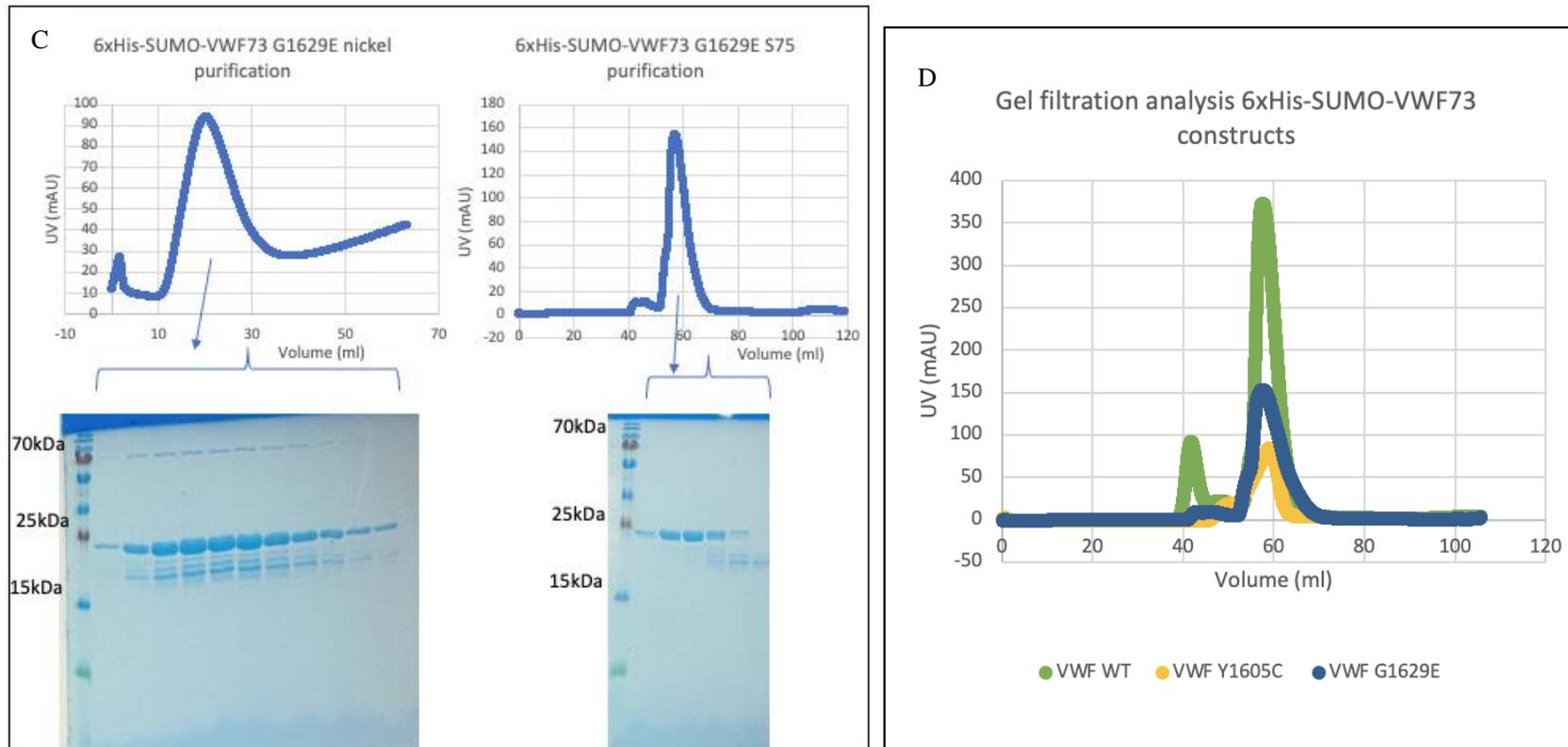
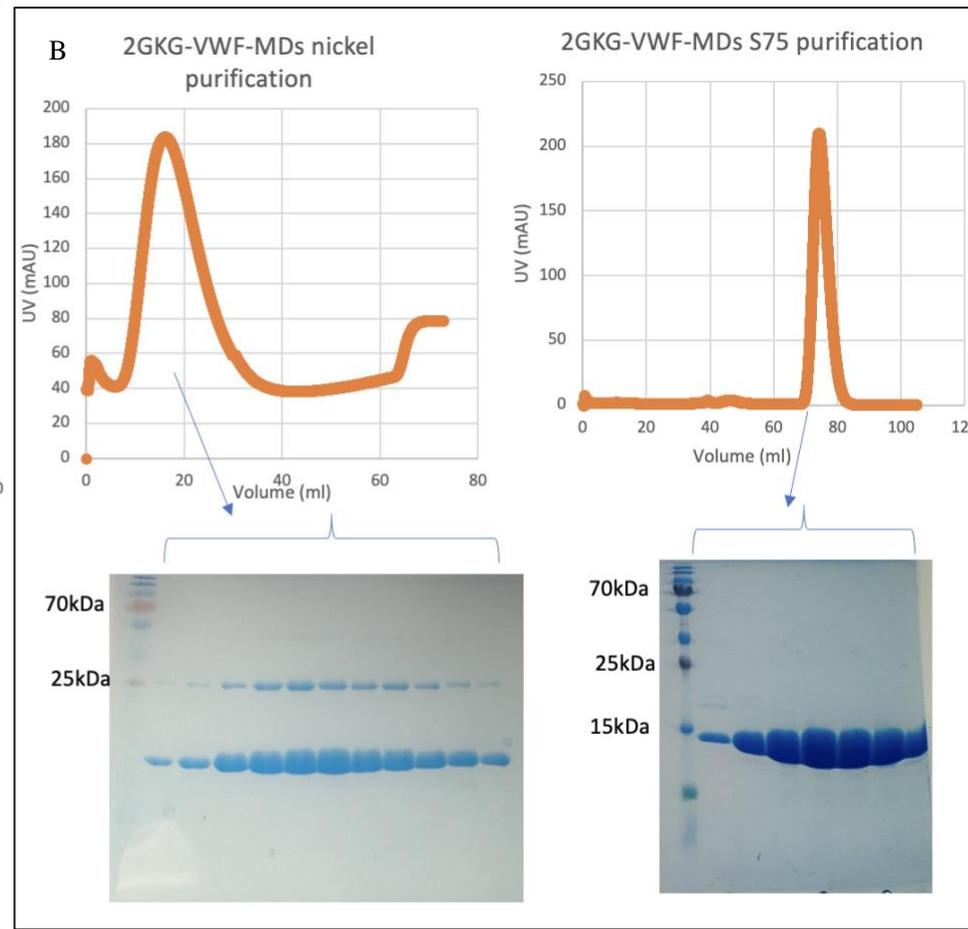
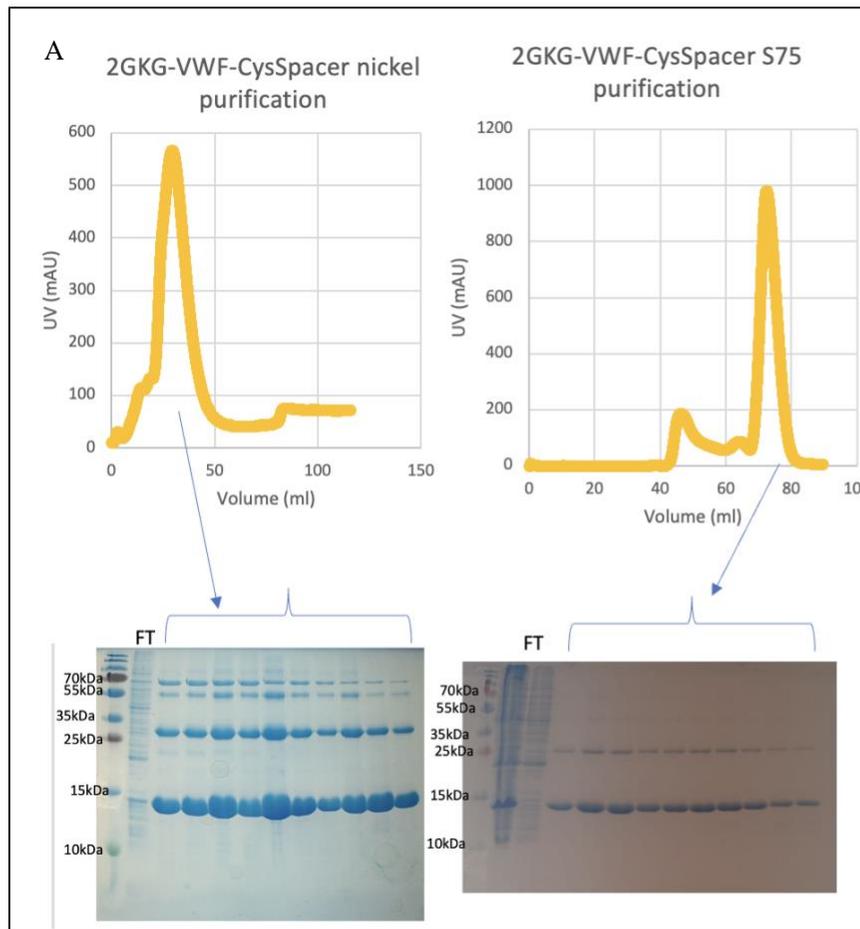
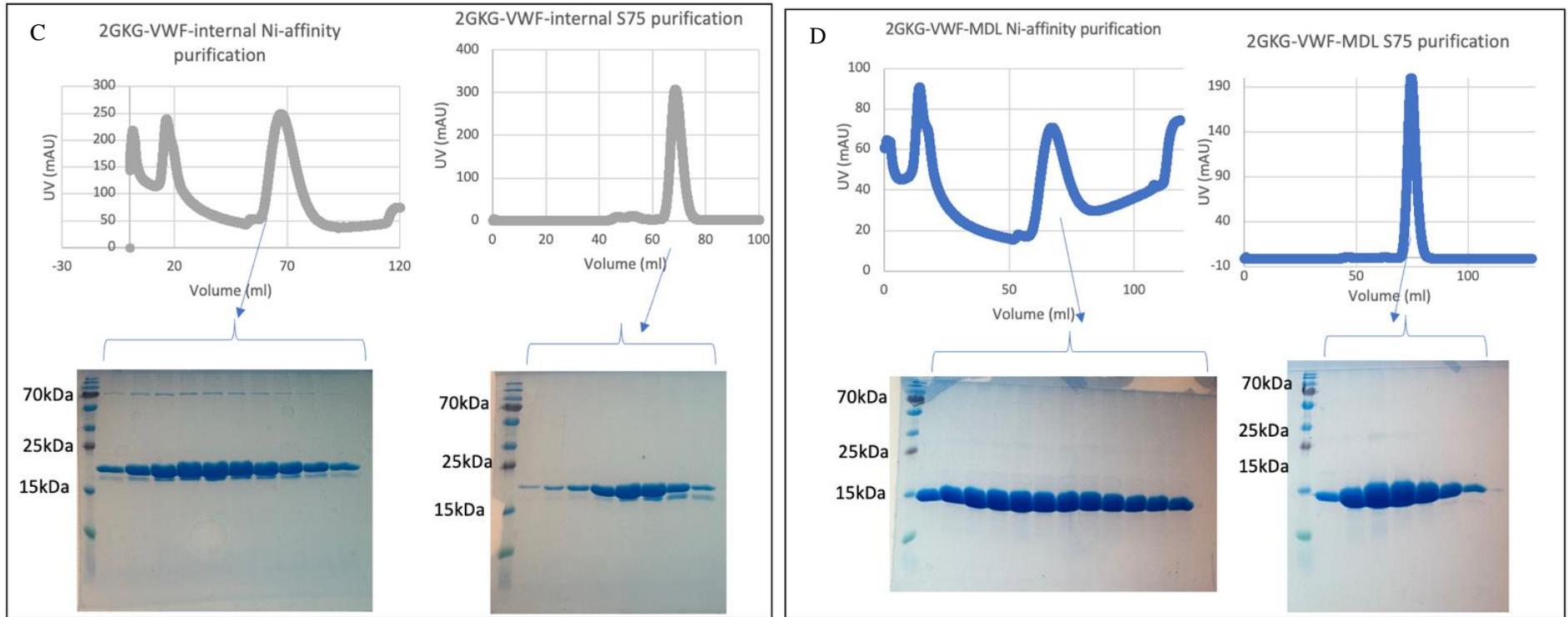


Figure 5.1: Purification of 6xHis-SUMO-VWF constructs. (A-C) Left chromatogram showing the elution profile of 6xHis-SUMO-VWF constructs from Ni-affinity column, protein containing fractions were confirmed by SDS-PAGE analysis. Right chromatogram showing the separation of the 6xHis-SUMO-VWF constructs on the S75 column. SDS-PAGE analysis of purification experiments fractions shown below respective chromatograms. (WT in green (A), Y1605C in yellow (B) and G1629E in blue(C)) the sample purity is higher as seen by the SDS-PAGE analysis. (D) Overlay of chromatograms from (A-C) with colours as in A-D.





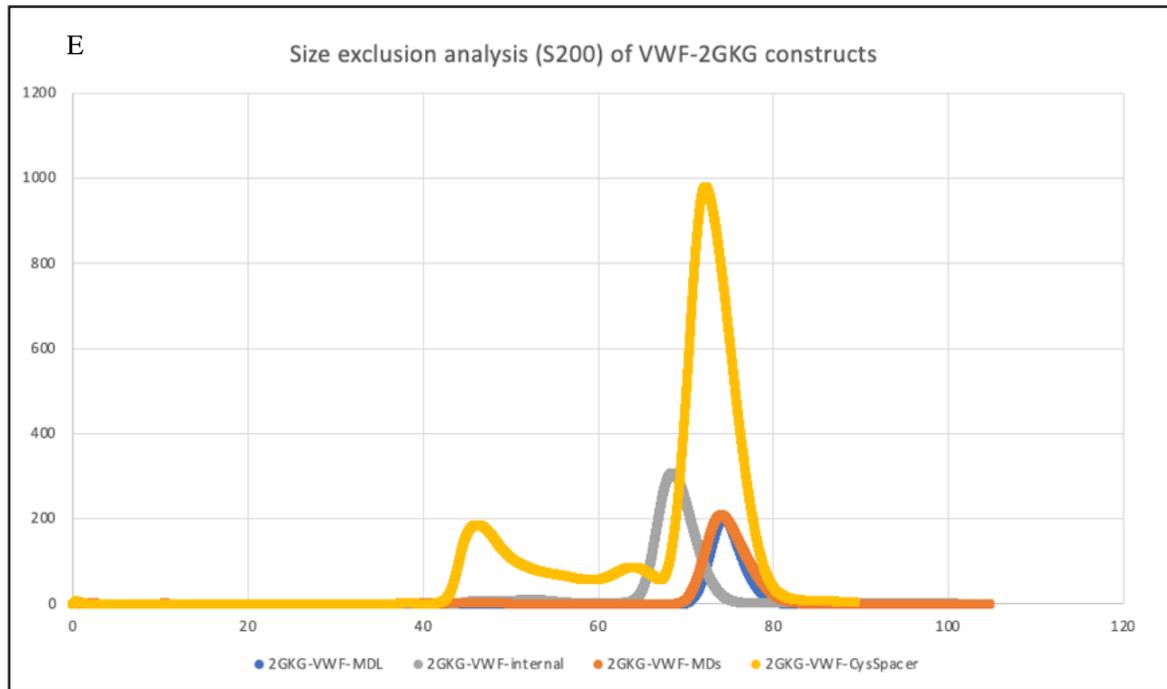


Figure 5.2: Purification of 2GKG-VWF constructs. Left chromatogram showing the elution profile of VWF-2GKG constructs. Right chromatogram showing the separation of the 2GKG-VWF constructs from the S75 column, SDS-PAGE analysis of purification experiments fractions shown below respective chromatograms (VWF-Cys-spacer in yellow (A), VWF-MDs in orange(B), VWF-internal in grey (C) and VWF-MDL in blue (D)). Protein containing fractions were confirmed by SDS-PAGE analysis, the sample purity is higher following S75 purification as seen by the SDS-PAGE analysis. (E) Overlay of chromatograms from S75 columns in A-D (colours as in A-D).

5.2.1.2 MDTCS E225Q mutant

The MDTCS construct was purified using the His-Excel column, with protein eluted at the 250mM imidazole step (Figure 5.3A). Protein containing fractions identified on the SDS-PAGE gel were then analysed using size exclusion gel filtration where the elution volume was ~75ml (Figure 5.3B). The SDS-PAGE gel indicated a clean purification however the protein concentration was lower than desired, so repeat expression and purifications were required. Western blot was also performed to confirm identity of MDTCS E225Q protein.

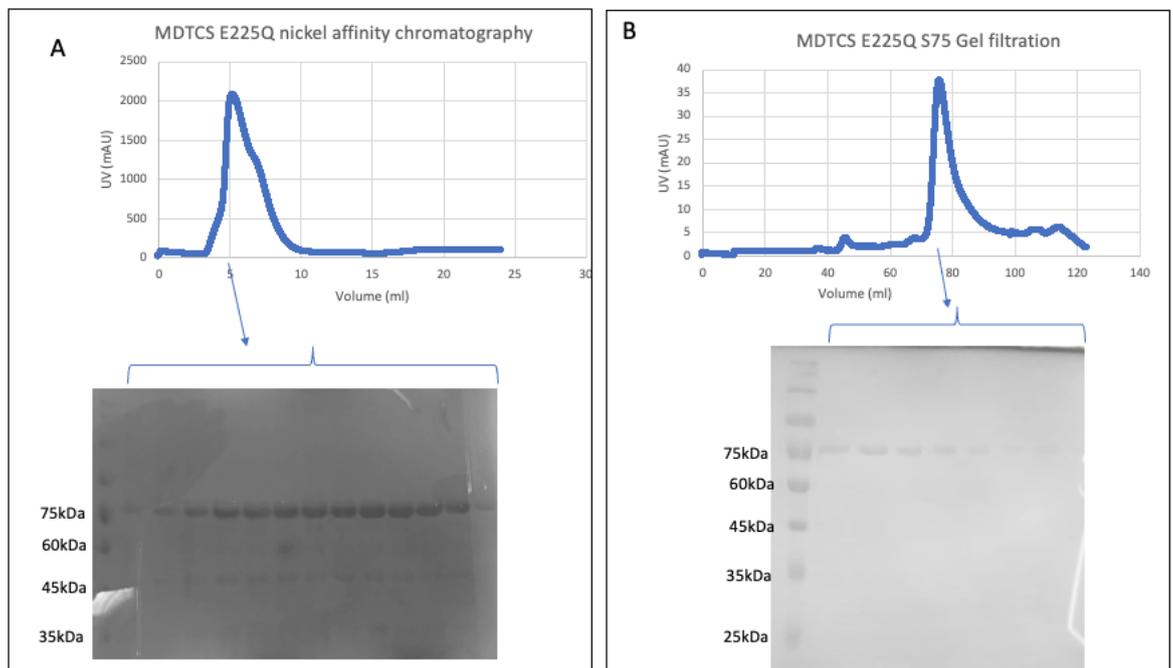


Figure 5.3: Purification of MDTCS E225Q construct. (A) Chromatogram showing the elution profile of MDTCS E225Q from Ni-affinity column, protein containing fractions were confirmed by SDS-PAGE analysis with gel shown below. (B) Chromatogram showing the separation of the MDTCS E225Q on the S200 column, protein containing fractions were confirmed by SDS-PAGE analysis with gel shown below. The sample purity is higher as seen by the SDS-PAGE analysis.

5.2.1.3 WT ADAMTS13 FL

A small-scale expression was carried out for WT ADAMTS13 full length (FL), and as activity was previously found to be highest in the media (unpublished data not shown) the protein was not purified, so media was frozen for use in activity assays.

5.2.1.4 DTCS-CUB

The DTCS-CUB construct was transiently expressed, and a small-scale expression and purification was performed. Purification involved both His-Excel and S200 gel filtration column. The elution volume and MW on SDS-PAGE did match that expected for a construct of 95kDa (Supplementary Figure 4)

5.2.2 Conformation of successful purification with western blot analysis and Mass spectrometry.

Western blotting was subsequently performed to confirm identity of the bands present. The Anti-6xHis antibody (Sigma) was used along with the secondary HRP-coupled anti-rabbit antibody (Sigma) for all constructs due to a lack of more suitable specific antibodies. The results confirm that all species contain this 6xHis-tag and are therefore suggests purification has been successful for all constructs (Figure 5.4). As well as this, MALDI-TOF mass spectrometry confirmed that the MDTCS E225Q and 6xHis-SUMO-VWF73 bands on the SDS-PAGE gels did contain the desired proteins, so purification was successful (Data not shown). However, analysis of the DTCS construct confirmed that the purified protein was an insect cell protein (Dmel\CG31705). More investigation into the viability to use this construct should be carried out to ensure expression of the protein is achieved, and how subsequent masking by the insect protein can be avoided. No further experiments were carried out on this construct due to funding and time restraints.

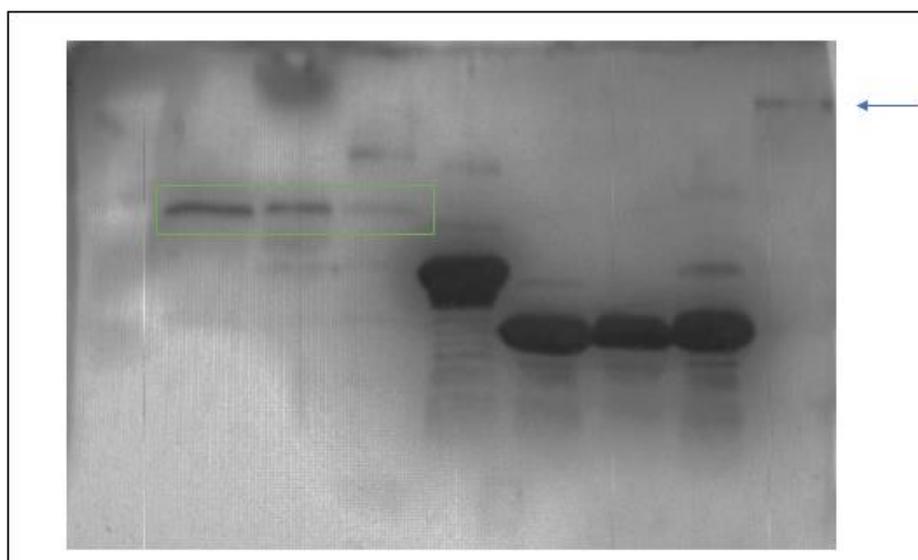


Figure 5.4: Western blotting confirming expression of poly-histidine tagged constructs. Loading from left to right contains: VWF constructs: WT 6xHis-SUMO,

Y1605C 6xHis-SUMO, G1629E 6xHis-SUMO, (contained in green box) 2GKG-internal, 2GKG Cys-Spacer, 2GKG MD-s and 2GKG MD-l and finally MDTCS E225Q (illustrated with arrow). Extra bands can be seen for higher molecular weight species of the VWF constructs.

5.2.3 Expression and purification of SUMO protease ULP1 for SUMO cleavage of 6xHis-SUMO-VWF73 constructs.

Expression and purification produced pure SUMO protease (Figure 5.5A) and aliquoted for use in the cleavage reactions for SUMO-tag removal of VWF constructs. The ULP-1 SUMO protease was incubated with VWF samples under standard conditions and repeats were carried out, however only partial cleavage was achieved as visualised in Figure 5.5B. This could be explained by the expression and purification conditions producing a protease with reduced activity, however there was differing levels of cleavage seen across the constructs with very minimal cleavage witnessed in WT 6xHis-SUMO-VWF. This issue coupled with difficult handling of the cleaved VWF peptide due its small size and flexible nature meant that the uncleaved 6xHis-SUMO-VWF was used in subsequent experiments.

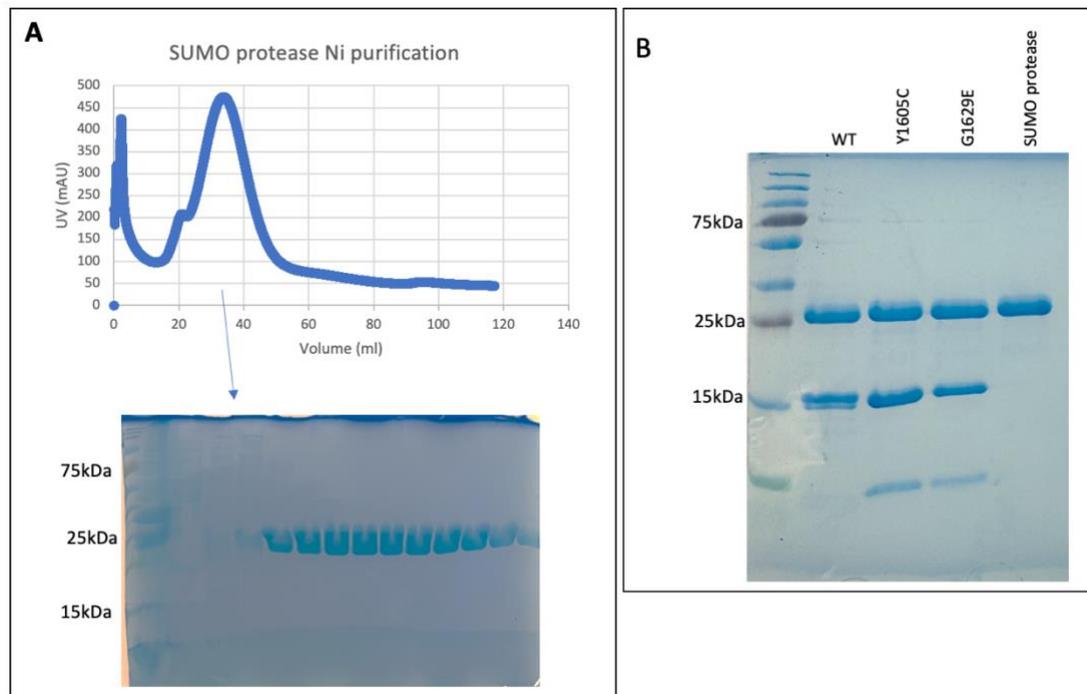


Figure 5.5 Purification of ULP-1 protein. (A) Gel filtration trace and SDS-PAGE gel of fractions from ULP-1 (SUMO protease) purification (B) SUMO protease (ULP1)

cleavage of 6xHis-SUMO-VWF73 constructs. The WT showed no cleavage and Y1605C and G1629E showed minimal cleavage of this SUMO tag. Control sample of ULP1 alone on the right.

5.2.4 ADAMTS13 and VWF cleavage assays

To further characterise the 6xHis-SUMO-VWF73 constructs, a cleavage assay utilising WT ADAMTS13 was used to check cleavage of 6xHis-SUMO-VWF73 WT and mutant constructs. As expected, cleavage was present in both WT and G1629E constructs (but not the scissile bond mutant Y1605C) as seen by the increase in cleavage products seen at 15 and 7kDa (Figure 5.6A). The 2GKG constructs do not show cleavage as expected when incubated with WT ADAMTS13 using the same conditions as the 6xHis-SUMO-VWF73 constructs (Figure 5.6B). This could be due to the large size of the 2GKG tag preventing access to the scissile bond on the internal, MD-S and MD-L constructs. No cleavage was expected on the Cys-spac construct. Further analysis of interaction is therefore to ensure the constructs are suitable for complexing with ADAMTS13 to support the docking information generated for design of the constructs.

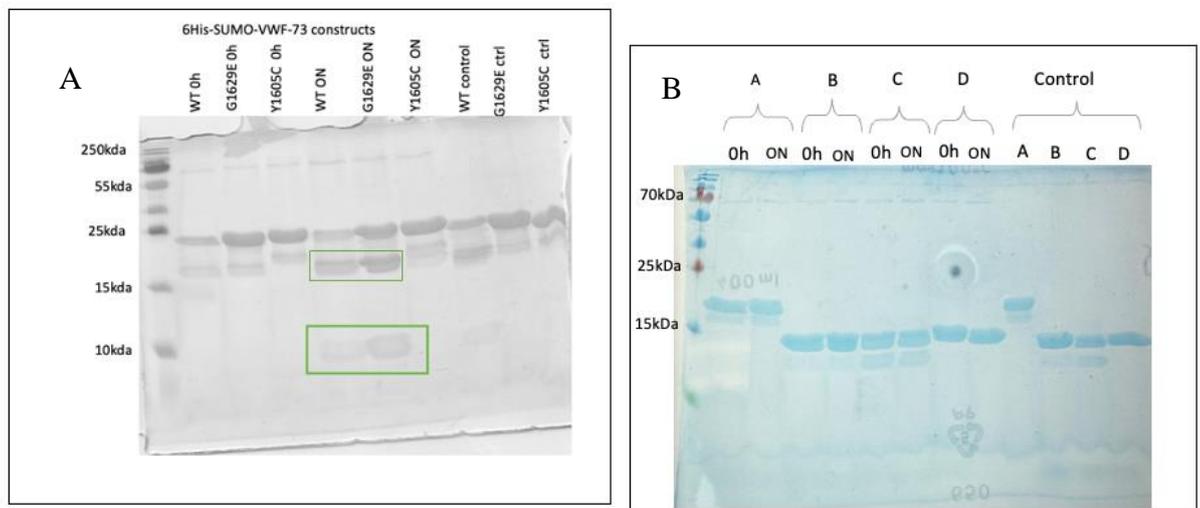


Figure 5.6 VWF constructs cleavage by ADAMTS13 (A) 6xHis-SUMO-VWF construct cleavage by WT ADAMTS13. Samples were taken at 0h, overnight (ON) and a control without ADAMTS13. Cleavage can be seen by a 10kDa product for WT and G1629E ON samples. (B) 2GKG-VWF construct cleavage by WT ADAMTS13. Samples were taken at 0h, overnight (ON) and a control without ADAMTS13. Cleavage could not be seen as no difference observed between control

and ON samples. Sample labels A= 2GKG-internal-VWF73, B= 2GKG-MP-L-VWF C= 2GKG-MP-S-VWF D=Cys-Spacer-2GKG-VWF.

5.2.5 Isothermal Calorimetry and Surface Plasmon Resonance for analysis of binding energetics between ADAMTS13 and VWF

Analysis of binding energetics was initially carried out utilising Isothermal Calorimetry (ITC) to provide a direct comparison with the WT data published by Petri et al (2019). The 6xHis-SUMO-VWF73 Y1605C mutant reported a K_D of 42nm with MDTCS E225Q suggesting stronger binding is occurring compared to the WT (Figure 5.7). However, it must be noted that more repeats are needed due to issues with background noise meaning many runs had to be abandoned, as well as the baseline drift that needs to be corrected.

To address the issues with high background from the ITC results, as well as the scarcity of MDTCS protein, SPR was chosen as the better route for understanding binding kinetics. The preparatory work for immobilising the MDTCS E225Q construct to the chip was carried out and experiments will be run by a postdoc in the lab for all VWF constructs. Preliminary results from these experiments reported a K_D of ~550uM which does not agree with previous ITC data published by Petri et al 2019, so further optimisation of this process may be required to understand this difference and generate meaningful results (Figure 5.8) (97).

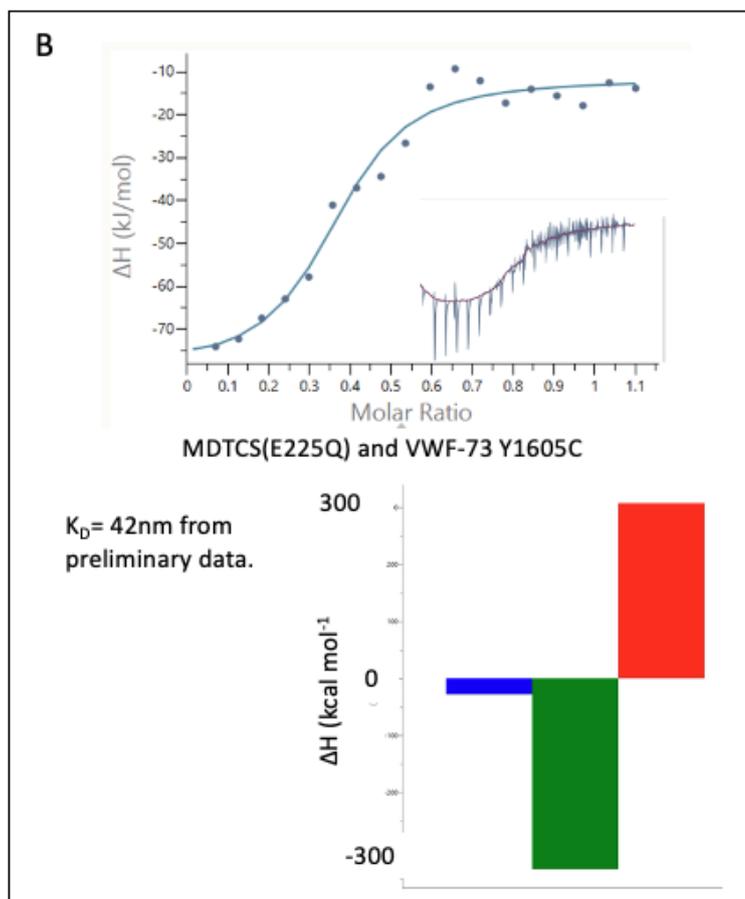
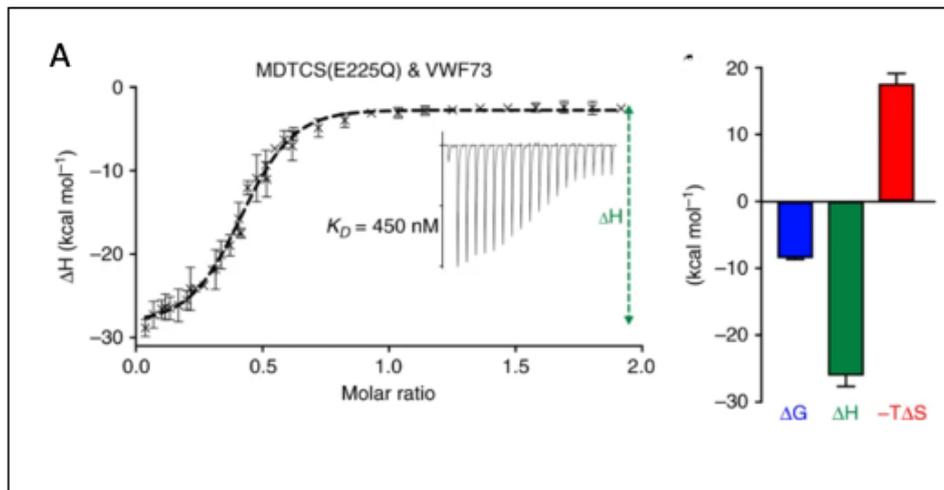


Figure 5.7 Binding analysis of MDTCS and VWF constructs. (A) Isothermal calorimetry (ITC) curves from Petri et al (2019) reported a K_D of 450nm for VWF-73 WT construct with MDTCS(E225Q). (B) For comparison, preliminary ITC data for the Y1605C mutant reported a K_D of 42nm suggesting stronger binding is occurring. More repeats are needed, and issues of background noise and baseline drift should be addressed.

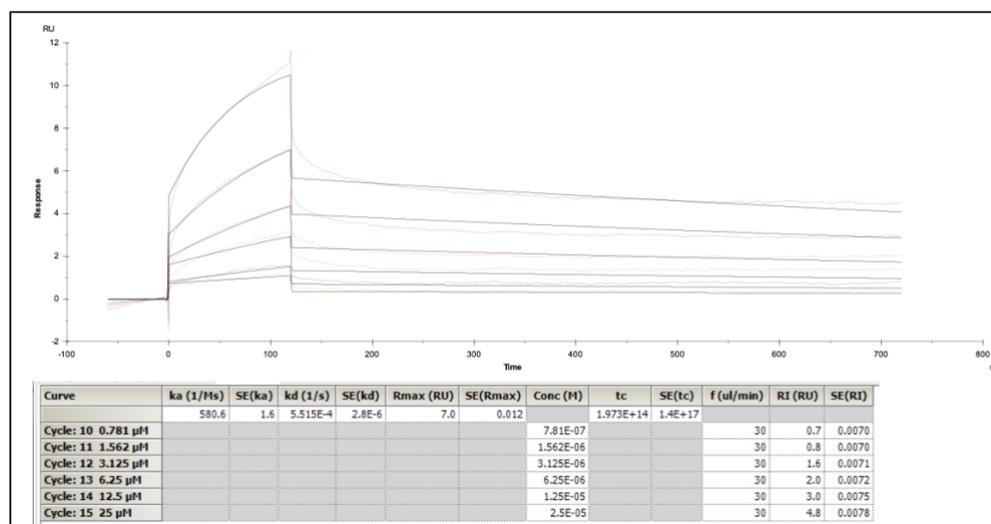


Figure 5.8: Binding curves and summary table from surface plasmon resonance analysis of MDTCS E225Q and 6xHis-SUMO-VWF73 WT. The K_D reported is 551.5um which disagrees with previously reported value by Petri et al. 2019 so more repeats are needed.

5.2.6 Complexing MDTCS with VWF

Analytical gel filtration was used to identify the interaction between MDTCS E225Q and 6xHis-SUMO-VWF73 constructs. As both proteins elute at similar volume on the HiLoad 16/60 Superdex 75 pg column, the aim was to investigate if the complex eluted at the expected volume for its MW (94kDa) or was affected by the unusual nature and elution volume shift of VWF constructs. Ratios of 1:1 and 2:1 were utilised for VWF constructs and MDTCS E225Q; and with each construct ran alone and then in complex on the Superose 6 to allow comparisons to be made.

Interestingly multiple peaks were observed and remained close in elution volumes, so clear margins were not achieved, however tips of the peaks could be separated to suggest an elution volume for the complex and therefore binding was taking place (Figure 5.9). The presence of multiple peaks of UV for the complex indicates that not all MDTCS has been bound. Surprisingly the complex of MDTCS and VWF exhibited a higher elution volume than either protein alone for VWF WT and Y1605C (Figure 5.9). There was not an observable shift seen in the peak for the MDTCS-G1269E complex so this was repeated utilising the S200 analytical column where a shift of elution volume could be seen. MDTCS-VWF WT complexes were also analysed on the analytical S200 for comparison, with a shift in elution volume was also present (Figure 5.9). By fractionating in small volumes (50ul), collection of fractions containing only the complex for use in crystallisation can be obtained.

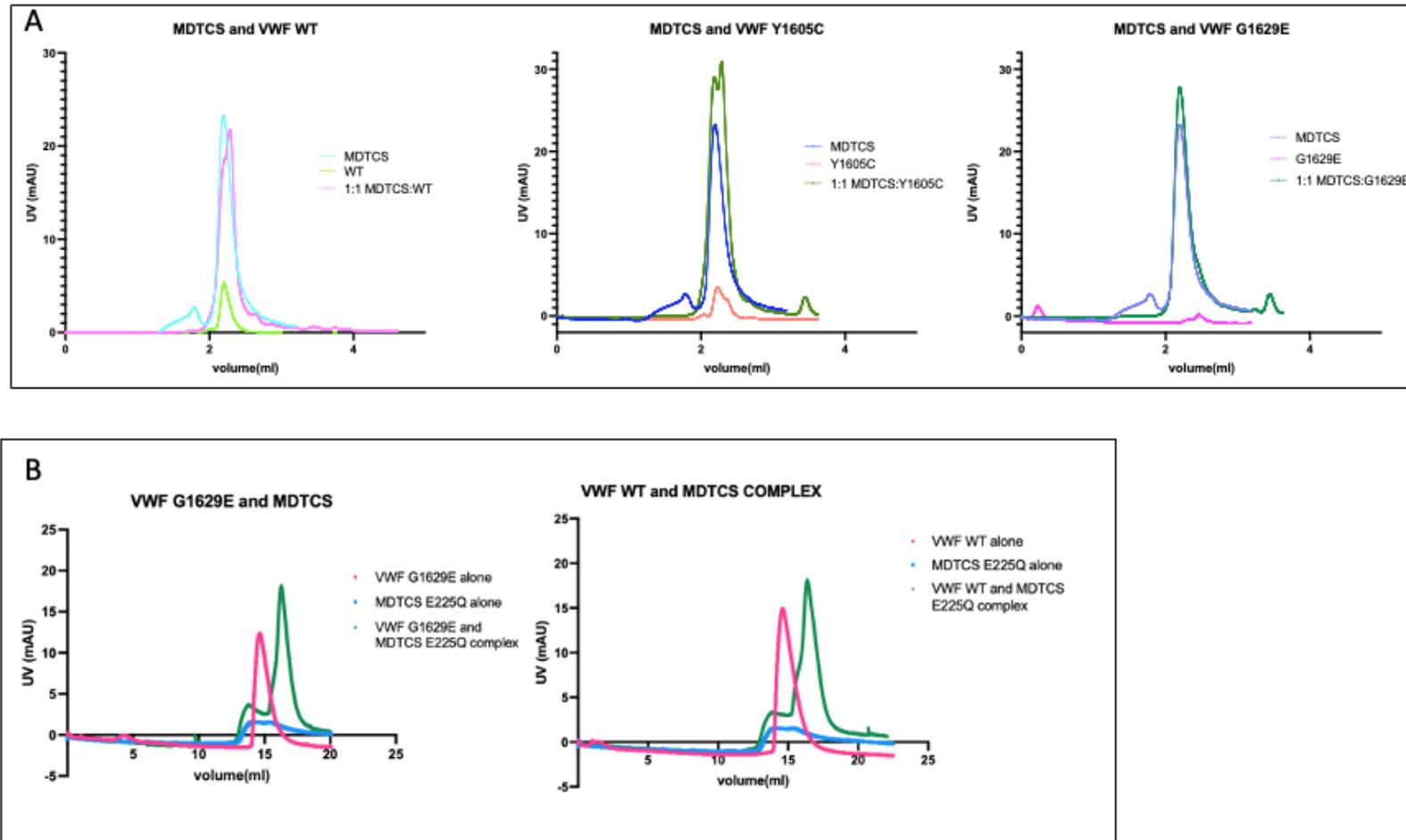


Figure 5.9 Size exclusion analysis to assess complex formation of MDTCS E225Q with 6xHis-SUMO-VWF73. (A) The Sepharose 6 column was utilised to analyse a 1:1 ratio of MDTCS:VWF constructs against the constructs alone. A shift in elution volume can be seen for VWF WT and

Y1605C (2.19ml MDTCS alone, 2.25ml Y1605C-VWF alone, 2.2ml VWF-WT alone and MDTCS-VWF complexes (WT and Y1605C) 2.3ml). Whereas VWF G1629E construct shows no indication of complex formation. (B) The S200 analytical column was utilised to repeat analysis for the G1629E VWF construct with MDTCS which did indicate complex formation as did WT VWF used for comparison (13.9ml MDTCS alone, 14.7ml G1629E-VWF and WT-VWF alone and MDTCS-VWF complexes (WT and G1629E) 16.4ml)

5.2.7 Crystallisation trials

Complexes obtained from ITC and size exclusion analysis were concentrated to 1-2mg/ml and used for crystallisation attempts. From the array of crystal trays set up from MDTCS E225Q and 6xHis-SUMO-VWF73 complexes, only a few crystals were observed (Figure 5.10). However, after visualising under a UV microscope it was confirmed these crystals were salt, likely from the calcium present in the purification buffer needed for MDTCS stability.

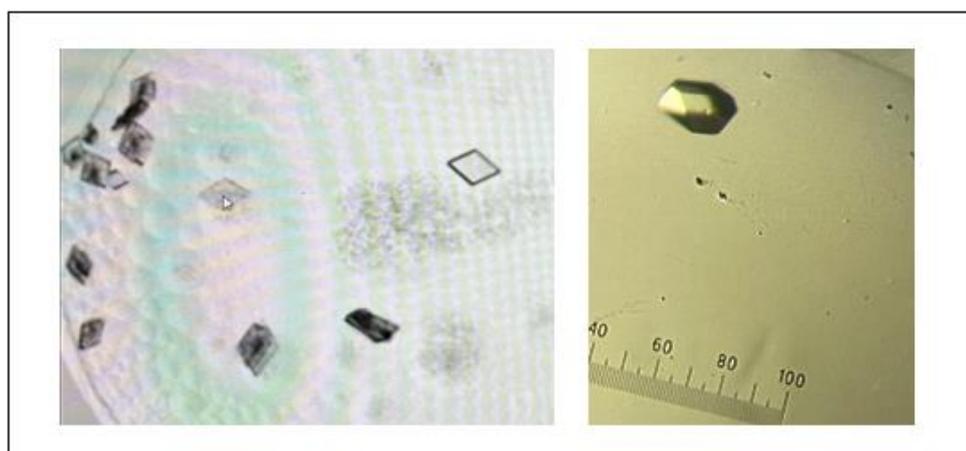


Figure 5.10 Crystallisation trials of MDTCS and 6xHis-SUMO VWF constructs. Images of MDTCS-6xHis-SUMO VWF WT complexes that formed crystals but did not produce diffraction.

5.3 Discussion

Following the design of constructs for crystallisation of ADAMTS13 in open and closed forms, it was important to characterise the new VWF constructs to support their utility in crystallography experiments.

The purification of VWF constructs suggests pure protein has been isolated from contaminants, however other species of VWF proteins are present. Due to unstructured nature of the VWF peptide its possible there is loss of the tag or peptide occurring to separate this from the 6xHis-SUMO tag. The western blot provides support for this with the His-Ab detecting both bands containing the 6xHis Tag. This purification product was taken forward to ITC and gel filtration analysis and it seems the other species do not interfere with the complexing of MDTCS and 6xHis-

SUMO-VWF so results can be interpreted with greater confidence. A similar issue can be seen in the SDS-PAGE gels of Del Amo-Maestro et al 2021 with multiple bands visible but no reports were made of issues with binding or cleavage by ADAMTS13(143).

The mutations Y1605 and G1629E were selected to improve binding to MDTCS E225Q to aid in complex formation and subsequent crystallisation. Of these mutants, Y1605C appears to show the greater improvement in binding and complex formation through co-ordination of the Zn^{2+} ion. There is evidence of the complex formation with size exclusion analysis and the ITC preliminary data with the K_D of 42nm compared to 450nm for WT VWF (97), supporting the concept for introducing the ionic interaction between the Zn^{2+} and cysteine. The YM1605-6CC VWF mutant was designed to further this interaction with multiple hydrogen bonds formed however the absence of expression may be caused by aberrant disulphide formation with two cysteines in close contact causing destabilisation of the construct. The G1629E mutant was designed to facilitate a confirmation suitable to binding ADAMTS13 as presented by Aponte-Santamaría et al (206). However, it may be that utilising the VWF G1629E mutant gives no improvement due to the already unstructured nature of the peptide. Utilising SPR binding analysis will confirm if these constructs are viable for crystallisation studies to find a VWF peptide that will have the required the binding affinity for MDTCS to form a stable complex suitable for crystallisation.

Size exclusion chromatography is an important method utilised to confirm the formation of the complex before setting up crystal trays or other structural biology techniques. When utilising size exclusion chromatography, it's expected the elution volume of a complex will decrease due to the increased molecular weight of a complex compared to proteins alone. However, in the analysis conducted on MDTCS and 6xHis-SUMO-VWF73 constructs the opposite, in fact, occurs (Figure 5.9). The increased elution volume may be explained by the unstructured nature of VWF constructs alone due to the flexible VWF73 peptide region of the construct, meaning they run at a lower elution volume than expected for a protein of their MW. This theory could be tested by running globular proteins in their denatured form to provide a more accurate standard curve (222). MDTCS also contains some regions

of flexibility including the metalloprotease domain and TSP-1, as shown by the use of the Fab for previous crystallisation of the MDTCS domains (97). Once in complex both the VWF and MDTCS constructs take a more structured complex conformation leading to the higher elution volume. When analysing complex formation between MDTCS and VWF, the presence of peaks for proteins alone suggests not all the MDTCS is bound by VWF. To address this problem and to have more of the complex available for crystallisation, the experiments can be scaled up with an increase in concentration of both proteins used, and an excess of VWF to ensure after incubation, all MDTCS is in complex with VWF.

The VWF-93 construct was previously utilised by Petri et al, along with the VWF peptide more recently also utilised by Del-Amo Maestro et al (97, 115, 143). The SUMO-tagged version of this VWF peptide has details published on purification, but no comments on the unstructured nature and it is not currently widely used in assays such as the FRET-tagged alternatives (223) The SUMO-tag cleavage of 6xHis-SUMO-VWF73 by ULP1 was not as efficient as reported previously resulting in a partially cleaved mixture of protein (Figure 5.5) (153). Even if cleavage was successful, the cleavage products of VWF peptides are particularly hard to work with due to the small size. For example, poor visualisation on SDS-PAGE gels or issues with a clear read of concentration on a spectrometer, which have also not been widely reported. For this reason, the uncleaved 6xHis-SUMO-VWF73 was utilised in the experiments rather than cleaved VWF. Complexes of 6xHis-SUMO-VWF73 constructs and MDTCS E225Q were set up for crystallisation, with the anticipation that the Y1605C mutant may improve crystallisation attempts with stronger binding witnessed. There are plentiful reasons why crystals do not form even with a plethora of screens and conditions trialled, many of which are yet to be elucidated. The SUMO-tag has not been investigated in detail as a crystallisation-tag and is often cleaved off before crystallisation, it was therefore reasonable to characterise new constructs utilising a tag designed specifically to aid crystallisation(224).

The 2GKG tag was designed and previously utilised to aid in USP-11 protein crystallisation as reported in Maurer, S. 2021 (208). The design of the 2GKG-VWF constructs allows targeting of specific areas of interest in ADAMTS13-VWF interactions, providing an alternative tag attempting to overcome the issues with the

SUMO-tag. A range of 2GKG-VWF constructs were utilised in attempt to explore the VWF-ADAMTS13 interaction. The 2GKG-Cys-spacer peptide is designed to bind DTCS, 2GKG-MD peptides (one with a short and one with a longer linker) to target MD domains and the internal 2GKG-VWF (replacing the TSP-1 domain) for interrogation of all exosites in MDTCS. These crystal structures would provide multiple observations crucial to understanding the binding and subsequent activation of both ADAMTS13 and VWF. It should be noted however the importance of gaining crystal structures for multiple 2GKG-VWF complexes to combine information on interactions of the domains to begin understanding of the interplay that occurs between domains, as seen in the differences between the published DTCS and MDTCS structures mentioned previously.

The expression of the 2GKG constructs was high as reported in previous use of the 2GKG tag, as was the presence of multiple higher weight species, detected in both SDS-PAGE and western blot analysis, suspected to be due to tag multimerization (208). As in the previous work this protein will be carried forward for further analysis into binding and complexing to confirm the suitability of these constructs, such as comparison to the K_D of WT and YC 6xHis-SUMO-73 with MDTCS. As well as this, investigation into the lack of cleavage by WT-ADAMTS13 should be explored. It's possible cleavage of the 2GKG constructs may produce cleavage products too similar in size to the uncleaved products so can't be visualised on the SDS-PAGE gel. Alternatively, the size or placement of the 2GKG tag may mean the scissile bond is inaccessible to ADAMTS13. Complex formation and binding analysis should be investigated as a priority to confirm these constructs are suitable for binding as the AF predictions suggest.

The expression and purification of the MDTCS E225Q construct largely followed the method reported in Petri et al 2019, with pure protein obtained following the purification method of Ni-affinity and gel filtration (97). However, due to the nature of the insect cell system compared to the bacteria system, the time required to optimise the expression took a considerably longer period of time with issues of contamination and low expression setting back experiments further still. Investigations into drosophila expression for the MDTCS construct would be worthwhile to ensure a high enough yield of pure protein is available to complete all

investigations into the VWF constructs and have sufficient protein available for setting up crystal trays.

The extensive glycosylation present on MDTCS may also hinder expression levels compared to other constructs which achieve higher yields in drosophila S2 expression system (225). To address these issues there are a few options to be explored. Initially a solubility tag could be used with the MDTCS construct for example the MBP-tag which has been successfully expressed and purified utilising drosophila system for crystallisation (226). The addition of this 42kDa tag may also help address the issues with overlapping elution volumes witnessed on gel filtration (Figure 5.9). In combination with this or as an alternative option, investigation into MDTCS glycosylation has revealed not all glycosylation sites are critical for successful expression and folding of MDTCS, so mutant MDTCS proteins could be explored to reduce the post-translational modifications present on MDTCS that may inhibit successful expression. This concept has been used successfully for other proteins as summarised in Chang et al 2007, with non-essential glycosylation sites mutated leading to successful protein crystallisation and structure solution (227). Finally, if sufficient funds or access to equipment was available, utilising equipment for improved S2 cell growth enabling monitoring of pH or air transfer has been successful in generating a higher protein yield; by reducing unfavourable conditions leading to excess cell death in the induction phase of S2 cell expression (228, 229).

The information obtained from the investigations into MDTCS and VWF constructs provide valuable insights into features of desirable constructs for ADAMTS13-VWF complex crystallisation. As well as this, improvements are suggested for future work exploring the interactions and activation mechanisms of ADAMTS13 and VWF.

6. Characterising VWF and novel nanobody ND6.

6.1 Introduction

Thrombotic Thrombocytopenic purpura, the rare life-threatening blood clotting disorder involving excessive blood clot formation in the microvasculature. Efficient diagnosis and quick treatment are required to minimise fatalities, and front-line treatment currently involves plasma exchange and corticosteroids. Consequently, there is a much-reduced fatality rate of approximately 10-20% however improvements can still be made to reduce this further (40, 59). Caplacizumab is the first anti-thrombotic FDA-approved drug for use in treatment for TTP (230). The mechanism involves stabilising VWF AIM-A1, thus requiring a higher shear force to achieve activation of and therefore increasing the threshold for mechanical activation of VWF(68). The AIM-A1 domain of VWF is an area of current research, specifically in the structural rearrangements that occur during activation. AIM-A1 refers to residues 1238-1493, with residues 1238–1271 (NAIM) and 1459–1478 (CAIM) termed the ‘autoinhibitory modules’ (AIM) flanking the A1 domain (231). Experiments thus far including hydrogen-deuterium exchange and single molecule force spectroscopy have been used to confirm the presence of the discontinuous AIMS, which stabilise the A1 domain, and termed a mechano-regulator of activity (68, 232). Stabilisation of the AIMS (for example by a nanobody) leads to a higher force requirement to activate A1, and therefore less force is required following disruption of the AIMS. Research into the structure and positioning of the AIMS and subsequently the interactions between the AIMS themselves and with the A1 domain is currently ongoing (68, 87).

Caplacizumab binds to and stabilises the NAIM and subsequently the A1 domain (68). The A1-caplacizumab crystal structure (PDB ID: 7A6O) provides a crucial insight into the mechanism by which the A1-nanobody complex leads to inhibition of GPIIb α binding to VWF. The rearrangement of α 1 β 2 loop and AIM residues when bound to caplacizumab provides steric hindrance to LBD of GPIIb α binding to A1 domain compared to unbound A1 (68). However, problems with increased risk of bleeding associated with caplacizumab treatment, albeit mild symptoms, means improvement to nanobody targeted VWF therapy can still be made (233). Novel nanobodies ND4 and ND6 are currently being explored as a new treatment to tackle

the problems presented by caplacizumab (88). Whilst the NAIM is targeted by Caplacizumab it is the CAIM that ND4 and 6 nanobodies bind to with K_D of 18.6nm and 604nm respectively (Figure 6.1) (234). ND4 and 6 were selected for crystallisation studies from the AIM-A1-enriched library of nanobodies due to their binding to the NAIM region of AIM-A1 and subsequently in a parallel-plate flow chamber experiment both nanobodies significantly inhibited VWF-mediated platelet adhesion to the collagen-coated surface under normal shear rates. ND4 and ND6 only partially inhibited platelet adhesion in whole blood at high shear rates, preserving VWF function at higher shear rates better than Caplacizumab (234). The allosteric regulation of GPIIb α binding, rather than direct inhibition, and partial inhibition of VWF activation at high shear may make the new nanobodies more suitable treatment options without the associated bleeding of caplacizumab treatment.

Due to the multiple crystal structures of A1 previously published (e.g.: PDB 7A6O, 1AUQ and 5VB8) there are multiple models that can be used to aid in the model building and refinement of data as well as advances in programs for generation of suitable nanobody models for use in structure solution. The complex crystal structure of A1-ND6 should reveal information about the interface with the CAIM and reveal structural rearrangements of AIM-A1 as well as potential information on glycosylation and activation states of VWF. N-glycosylation of VWF is found across many domains of VWF and is well characterised, whilst o-glycosylation was previously less so (78). O-glycosylation is however clustered mostly around the A1 domain on the N/CAIMs(235). The structure of these O-glycans have been revealed as heterogenous consisting of sialiated core 1 and 2 structures, but their role in VWF structure was previously unclear (236). However recent papers have suggested an importance for the O-glycans in altering flexibility of the N/CAIM and subsequent exposure of the A1 domain for GPIIb α (77). The mechanism by which this occurs has not been fully resolved, so more structural information on the A1 domain and O-glycosylation would therefore be beneficial to give a more complete picture on VWF activation and GPIIb α binding.

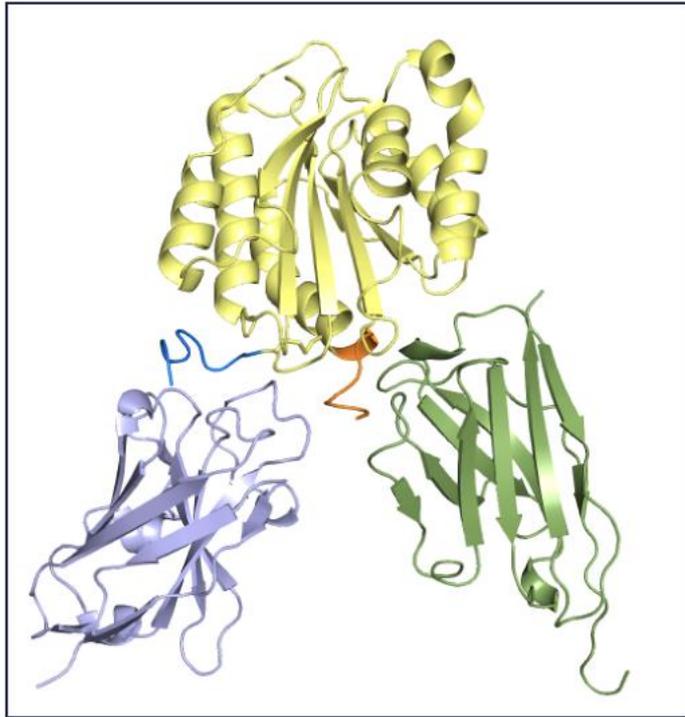


Figure 6.1: Cartoon of AIM-A1 structure with Caplacizumab and ND6 localised to binding sites. Cartoon diagram of the crystal structure of the vWF AIM-A1 domain (yellow) in complex with ND6 (green) and caplacizumab (purple) (PDB: 7A6O) with the CAIM in orange and NAIM in blue. The binding sites of the nanobodies are distinct but with both targeting the AIMS of AIM-A1. (ND6 structure unpublished on the PDB but presented in this chapter).

6.2 Results

6.2.1 X-ray diffraction experiments

Shard-like crystals of VWF AIM-A1/ND6 were observed in a complex Morpheus III mix (1.2 % Cholic acid derivatives mix, 0.1 M Buffer System 3 pH 8.5 and 30 % Precipitant Mix 1) (Figure 6.2A). The single crystal shards grew from dense blobs of protein aggregation but could be fished out to largely separate the crystals from these non-crystalline blob structures thus were suitable for X-ray diffraction experiments. Multiple crystals were subjected to X-ray diffraction due to the thin nature of the shards rendering them susceptible to breakage, a line scan was also used in an attempt to overcome issues with radiation damage by spreading the dose of radiation. Analysis of the diffraction images obtained however indicated issues with the data on multiple crystals including anisotropy and radiation damage towards the end of the images of certain crystals so interrogation of images would be required to obtain

data not limited by these issues (Figure 6.2B-C). Crystallisation statistics can be found in Table 6.1.

Table 6.1: Crystallographic data collection and refinement statistics.

	VWF AIM-A1 ND6	ND4
Data collection		
Space group	P 21	C 1 2 1
Cell dimensions		
a, b, c (Å)	38.8 174.3 197.3	35.35 63.19 50.57
α, β, γ (°)	90.00 89.96 90.00	90 94.97 90
Resolution (Å)*	199.991 - 3.343 (3.7 - 3.3)	27.2 - 1.15 (1.19 - 1.15)
R_{merge}^a	0.263 (0.808)	0.121 (0.905)
$CC^{1/2}$	0.981 (0.389)	0.998 (0.485)
$I / \sigma I$	3.7 (1.7)	11.30 (0.85)
Completeness (%)	83.1 (48.5)	81.64 (25.92)
Multiplicity	3.3 (3.4)	5.9 (2.4)
Wavelength	0.999	0.95374
Refinement		
No. Reflections	70134 (3611)	187864 (2407)
R_{work}^{b*}	0.243	0.1554
R_{free}^*	0.274	0.1619
Average B factor	95.7	14.4
R.m.s. deviations		
Bond lengths (Å)	0.007	0.014
Bond angles (°)	1.8	1.76
Ramachandran statistics ^d		
Ramachandran favoured (%)	91.88	97.62
Ramachandran allowed (%)	7.81	1.59
Ramachandran outliers (%)	0.31	0.79
Rotamer outliers (%)	0.54	0

*Values in parentheses are for highest-resolution shell.

^a $R_{\text{merge}} = \text{Sum}(h) [\text{Sum}(j) [I(hj) - \langle Ih \rangle] / \text{Sum}(hj) \langle Ih \rangle]$ where I is the observed intensity and $\langle Ih \rangle$ is the average intensity of multiple observations from symmetry-related reflections calculated.

^b $R_{\text{work}} = \text{Sum}(h) ||F_o|/h - |F_c|/h| / \text{Sum}(h)|F_o|/h$, where F_o and F_c are the observed and calculated structure factors, respectively. R_{free} computed as in R_{work} , but only for (5%) randomly selected reflections, which were omitted in refinement, calculated using REFMAC.

^c $CC^{1/2}$ = Pearson correlation coefficient,

^d Ramachandran statistics were calculated using Molprobity.

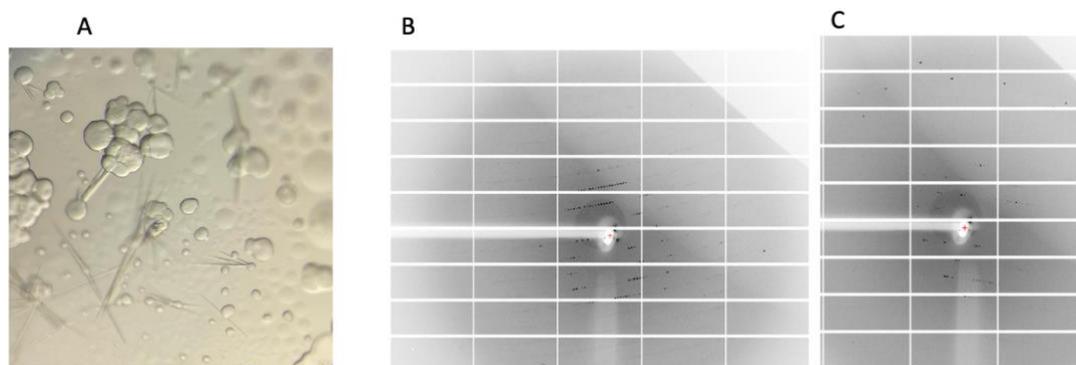


Figure 6.2: Images of crystals and their diffraction of AIM-A1/ND6 complex. (A) Crystal formation of A1-ND6 complex grown in 1.2 % Cholic acid derivatives mix, 0.1 M Buffer System 3 pH 8.5, 30 % Precipitant Mix 1. Thin shards growing out of a dense crystal collection forming a sphere. (B) and (C) Diffraction images of an A1-ND6 crystal collected by a line scan. Images show resolution is 3.4Å with pathologies of anisotropy, and radiation decay (C) to be addressed in processing.

6.2.2 Crystal packing and space group determination

The output of automatic data processing in IsPyB gave multiple space group suggestions of P 2 2 21 and P 1 21 1 for the diffracted crystals. Initially processing in the orthorhombic space group (P2 2 21) seemed successful, after anisotropy correction however R-factors did not reach above 0.4. After closer inspection of the data, the P 1 21 1 space group was selected containing a very slight beta angle of 89.954°. This beta angle shift from 90° was not detected by all processing software, explaining the previous orthorhombic space group selection with all angles at 90°. Processing in the monoclinic space group provided R-factors around 0.3 before refinement giving confidence for this solution.

6.2.3 Differences between subunits in the asymmetric unit (ASU)

Using Matthews coefficient calculation in CCP4 it was indicated that 4 copies of both A1 and ND6 were present in the ASU (Figure 6.3). Within the ASU not all complexes share the same environment, and this explains the small differences seen between the chains. It's important to note due to the flexibility around the solvent channels certain regions of both A1 and the nanobody show differences between the chains, so considering all chains can be important to find the best region of density at the area of interest for confidence in results. The VWF A1 domain is resolved to a

different degree in each subunit, with clearest density witnessed in chain B with Y1258 in the proximal NAIM visualised and in chain A with A1474 in the distal CAIM visualised. The lack of stabilising structures around the flexible N-terminal and C-terminal regions can explain differences in determination of the structure.

Central to all four subunits of the nanobody is a region containing cholic acid derivatives from the crystallisation condition with each nanobody chain containing a binding pocket for the central sterol core of the cholic acid derivatives (Figure 6.4). Both cholic acid (CHD) and CHAPS (CPS) molecules were modelled into the extra density between the nanobody chains, the density is not clear enough to determine the location of all CHAPS atoms, but it was modelled where best fitted the density. The interaction between ND6 and CPS interactions were best visualised between CPS in chain D and ND6 chain H, with a hydrogen bond present between S118 and weaker interactions with L12, P120, S117, P116 and V116. (Figure 6.4B).

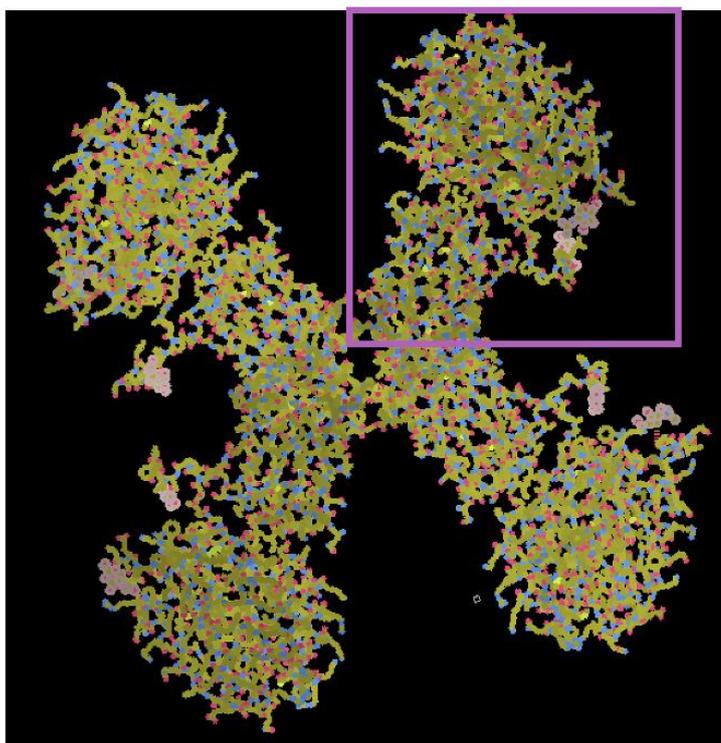


Figure 6.3: ASU of AIM-A1/ND6 crystal. Four copies of VWF AIM-A1 and ND6 present within the ASU, with a solvent channel present in the centre. One copy of the AIM-A1 ND6 complex is shown by the purple box.

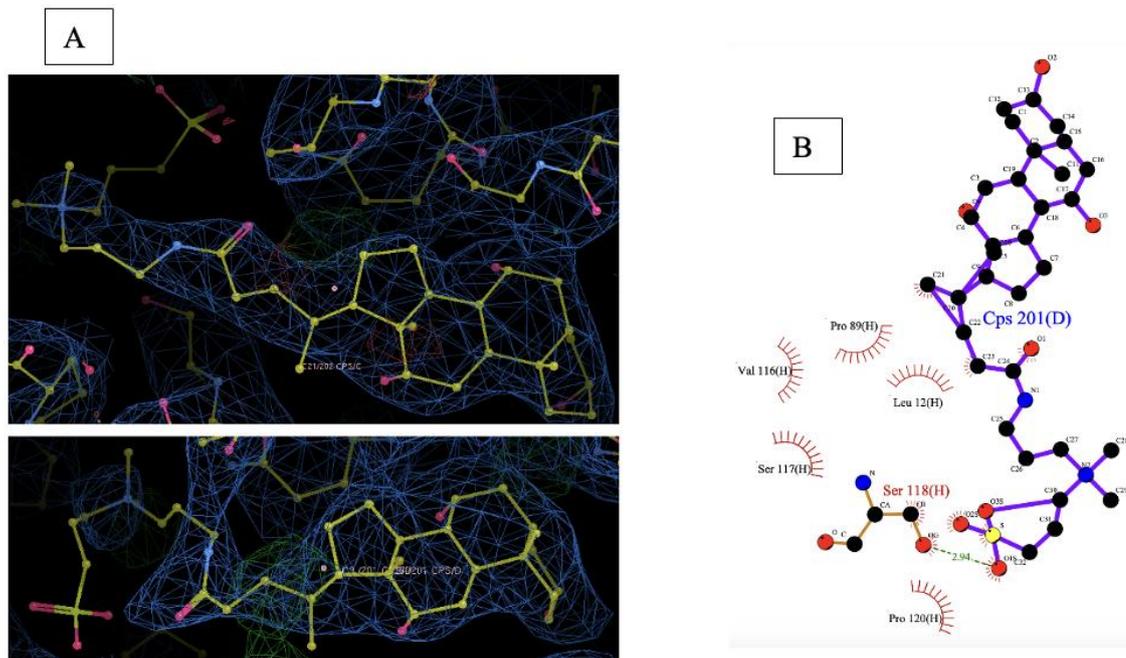


Figure 6.4 CPS molecules facilitate the AIM-A1/ND6 crystal formation. (A) Electron density surrounding the CHAPS (CPS) molecules present in the solvent channel between ND6 chains. (B) The PDBSUM server (utilising LigPlot) created a schematic of the interaction between CPS (chain D) (on the right) and ND6 chain H (Illustrated with residue names on the left) (149). Hydrogen bonds are shown as green dashes and van der Waals interactions as red dashes around ND6 residues.

6.2.4 A1-ND6 interface interactions

To characterise the interaction between A1 and ND6, the crystal structure of this complex was resolved to 3.4 Å (Table 6.1). ND6 comprises antibody recognition loops (or Complementarity-determining region) CDR1, CDR2, and CDR3 of 9, 15, and 10 residues respectively. The sequences and locations of each CDR loop are depicted in Figure 6.5. Specifically, I29 and Y33 in CDR1 of ND6 form hydrogen bonds with E1434 of vWF A1 domain, and D78 in the constant region of the nanobody makes a salt bridge with vWF residue K1408. There is also a hydrogen bond interaction with the CAIM formed by S55 and S57 in the CDR2 and A1464 of vWF (Figure 6.6). The rest of the interactions between CAIM and A1 form only weak Van der Waals forces which is expected due to the relatively weak binding observed between AIM-A1 and ND6 (234). Stabilisation of the AIMS, provided by ND6, also allowed a novel section of the AIM-A1 (residues 1467-1471) to be resolved for the first time. This includes the ristocetin binding site found at proline rich section of the CAIM (Glu1463–Asp1472) (Figure 6.7).

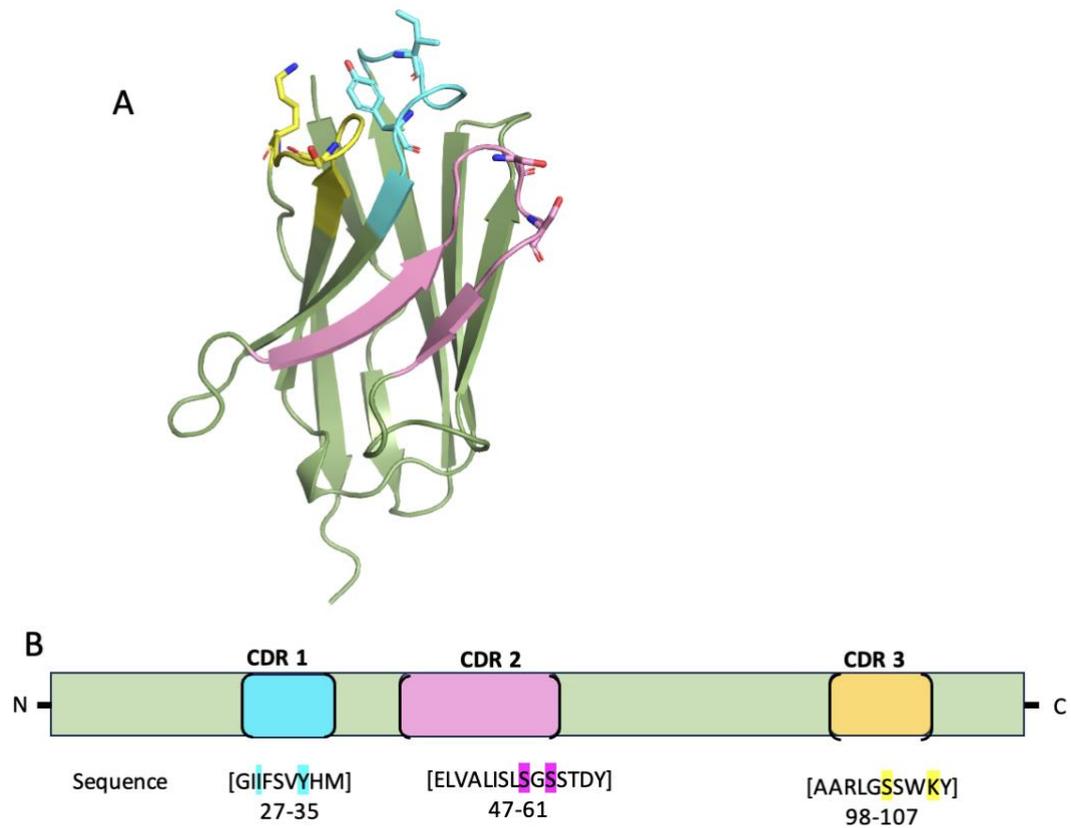


Figure 6.5 ND6 nanobody cartoon and schematic representation. A) Cartoon representation of ND6 (green) highlighting CDR1, 2 and 3 in cyan, magenta, and yellow, respectively. Amino acids engaged in binding with vWF are depicted as sticks. B) Schematic diagram of specific sequences and locations of CDR loops in ND6. Residues involved in binding to VWF AIM-A1 are highlighted in the sequence.

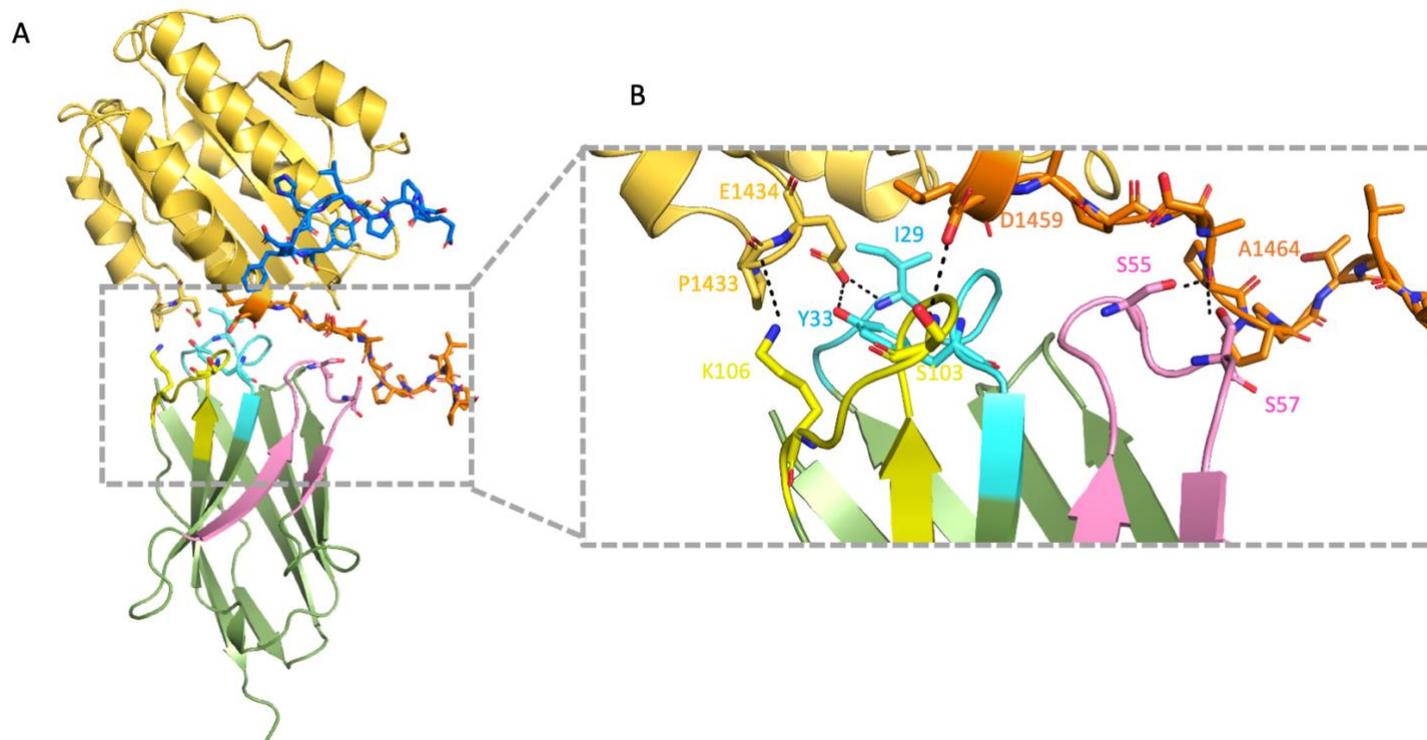


Figure 6.6: VWF AIM-A1 and ND6 complex interface. A) Cartoon diagram of the crystal structure of ND6 (green) in complex with the vWF AIM-A1 domain (yellow), with CDR1, 2, and 3 colored cyan, magenta, and yellow, respectively. The NAIM region of AIM-A1 is shown in blue and CAIM in orange. B) A close-up view of the interface interactions. Specifically, ND6 Ile29 main chain nitrogen and ND6 Y33 side chain hydroxyl form hydrogen bonds with Glu1434 of VWF, and the ND6 His34 sidechain hydrogen bonds with the main chain of Asp1459 of VWF. In addition, there is a hydrogen bond interaction formed by ND6 Ser55 and Ser57 side chain hydroxyls and the Ala1464 main chain carbonyl of VWF, and a salt bridge formed by Asp78 of ND6 and Lys1408 of VWF.

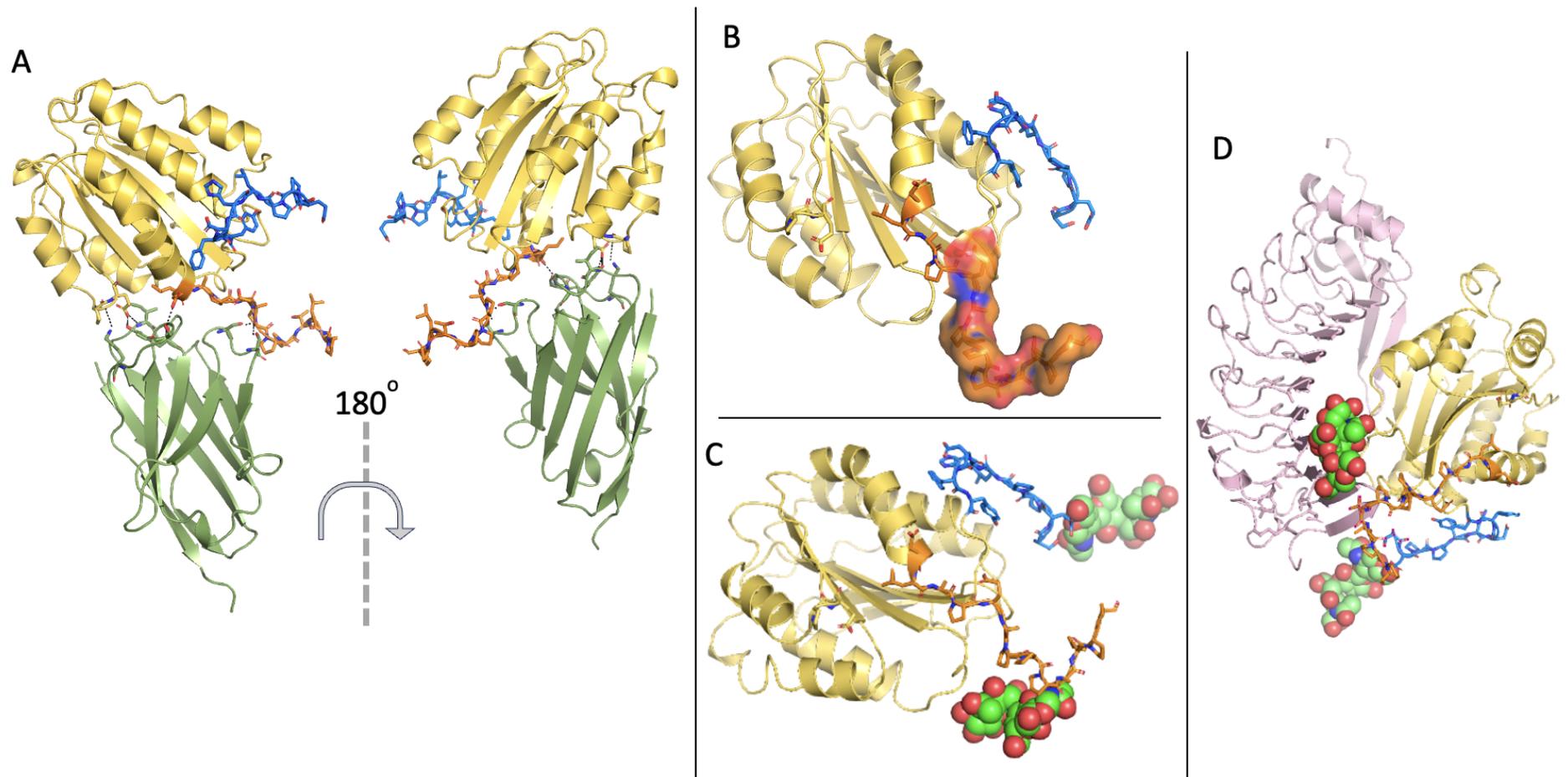


Figure 6.7 VWF AIM-A1 and ND6 cartoon and molecular models of experimental structure. (A) Cartoon representation of AIM-A1 ND6 complex (ND6 in green and vWF AIM-A1 domain in yellow), with the CAIM (orange) and NAIM (blue) extending away from the A1 domain. (B)

Cartoon representation of AIM-A1 showing the location of the ristocetin binding site (orange mesh) on the CAIM, colours as seen in (A). (C) Cartoon representation of AIM-A1 with O-glycans (green spheres) on T1468 and S1263. O-glycans consist of A2G and SIA monomers, enabling visualisation of their location in relation to AIM-A1. (D) Cartoon representation of AIM-A1 (colours as in (A) and (C)) with GPIIb LBD (PDB code: 1SQ0). Severe clashes are observed between the LBD and CAIM and O-glycan residues when structures are overlaid.

6.2.5 Structural differences of ND6 bound AIM-A1, to unbound AIM-A1 and caplacizumab bound AIM-A1 structures.

The CAIM adopts a conformation different to that seen in unbound A1 (PDB: 1AUQ (237)), specifically past E1463. Similarly, there is also structural rearrangement of the NAIM with both AIMS extending away from the main body of the A1 domain (Figure 6.7). Furthermore, some differences can be seen in specific regions around A1 accommodating this structural rearrangement (Figure 6.8). The beta sheets B and C shift conformation in AIM-A1 ND6 bound, (D1323 loses the hydrogen bond to S1356 and instead interacts with S1324 and H1322 loses hydrogen bond to S1356) allowing movement of the NAIM. The alpha helix region 5 exhibits loss of the hydrogen bond between S1394 and F1397 as well as R1392 to R1395. The alpha helix 7 region (V1443 to L1457) maintains the alpha helix shape but its position relative to the centre of the A1 domain is altered with hydrogen bond interaction changes compared to unbound A1. Specifically, interactions from Q1449 to I1453 and L1446, R1450 to V1454 and L1446 as well as Q1448 to E1445 are no longer seen in the complex, whilst novel hydrogen bonds are seen from Q1449 to E1445 and E1452. This change in alpha-helix positioning allows for the movement of the CAIM and accommodates binding to ND6 compared to unbound A1 (Figure 6.8).

Comparison of the AIM-A1 structure when bound to ND6 or caplacizumab (PDB: 7A6O), exhibits a difference in the Beta 2 and 3 sheets as seen with 1AUQ as well as all changes in the a5 loop. The main difference is the location of the AIM residues, and its notable the AIM-A1 ND6 structure has more resolved terminal CAIM residues. In the ND6-bound structure, these would be found in the space occupied by LBD of GPIIb α when structures are overlaid (Figure 6.7)

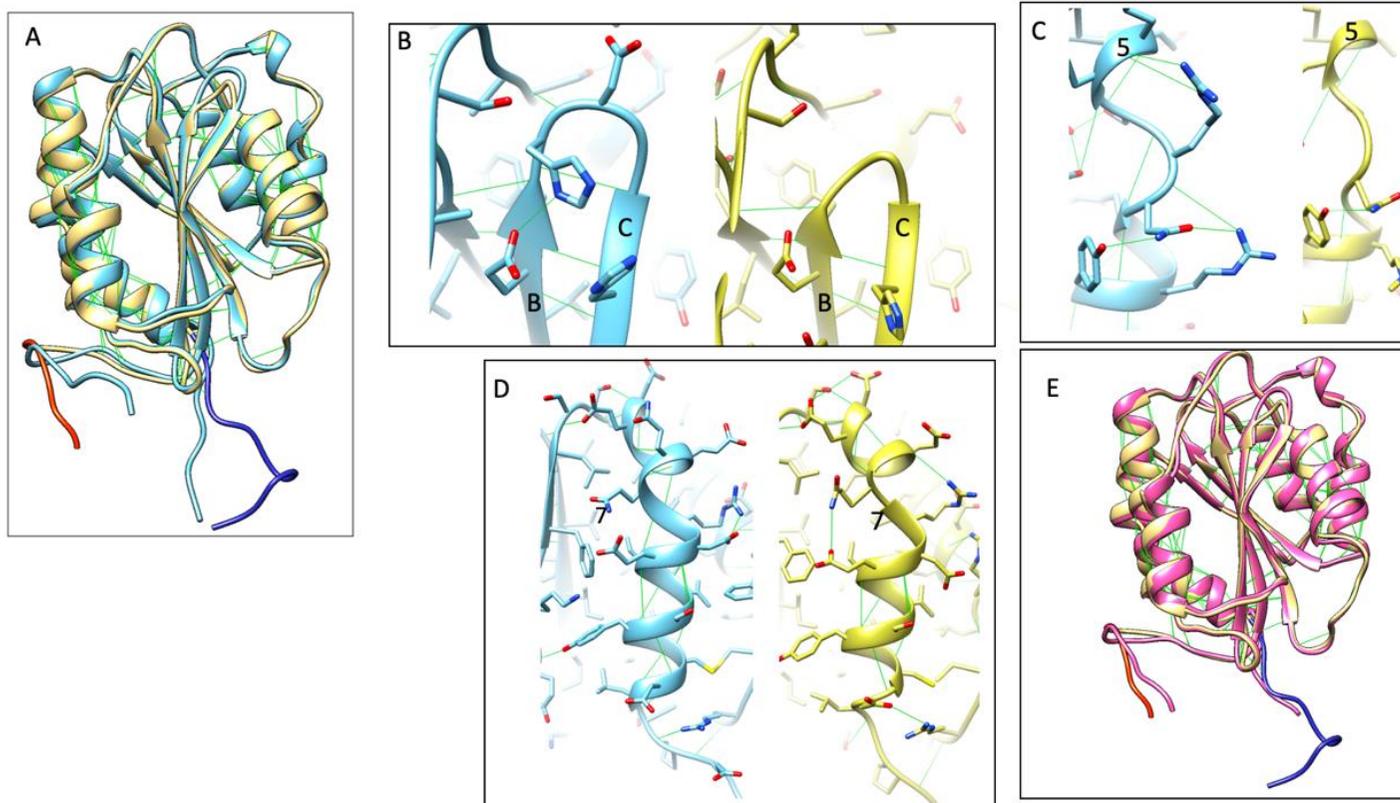


Figure 6.8: (A) Cartoon representation of VWF AIM-A1 in yellow (CAIM in orange and NAIM in dark blue) and AIM-A1 from 1AUQ in light blue. Hydrogen bonds depicted as green lines, highlighting area with hydrogen bond differences and changes in secondary structures. Close ups of beta sheets B and C, and alpha helix regions 5, and 7 can be seen in (B) (C) and (D) respectively. (E) Cartoon representation of AIM-A1 in yellow (CAIM in orange and NAIM in dark blue) and AIM-A1 from 7A6O in pink. Hydrogen bonds depicted as green lines illustrating similarities in secondary structure with some hydrogen bond changes.

6.2.6 O-glycosylation of AIM-A1

O-glycosylation appears to be present at residues T1468 and S1263 which haven't been resolved in previous structures (Figure 6.7C-D and Figure 6.9). The density around this area does not clearly define the carbohydrate molecules however it's recently been elucidated that mono-sialylation accounts for 78% of O-glycosylation and 22% are di-sialyated in AIM-A1 domain (238). The small molecule A2G was inserted into areas of electron density surrounding both T1468 and S1263 to give an indication of localisation and space taken up by these O-glycosylation chains. The O-glycans also play a large part in the space occupied by the CAIM, and the intrusion this causes when overlaid with A1- GPIb α (Figure 6.7D).

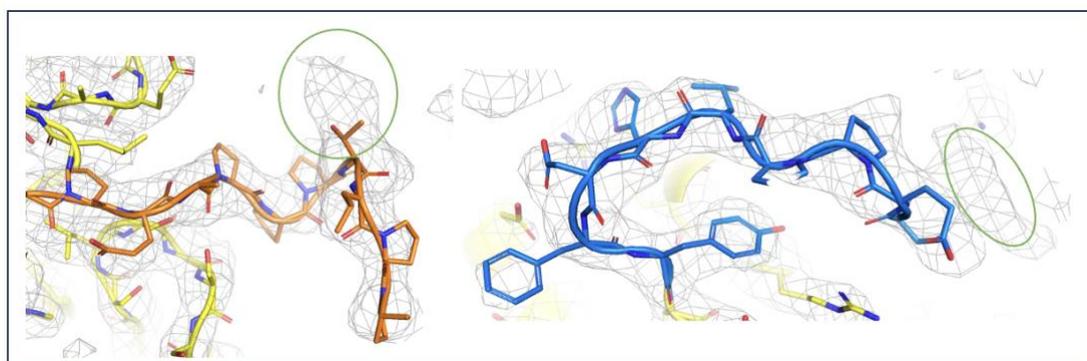


Figure 6.9: AIM structure with electron density map. Molecular visualisation of the crystal structure of the vWF AIM-A1 domain (yellow) with CAIM in orange and NAIM in blue. A green circle highlights the area of density around T1468 and S1263 where the O-glycosylation is expected.

6.2.7 ND4 high resolution structure

The ND4 crystal structure was resolved to a high resolution of 1.15Å, much greater detail can be seen in positioning of the side chains of amino acids, the clearly defined electron density can be seen in Figure 6.10A. The data collection statistics after data reduction were good quality, with $CC_{1/2}$ of 0.998 and completeness of 81.64 (Table 6.1). These statistics could be improved by combining multiple crystals however the gain in data would not be beneficial due to the presence of the nanobody alone.

The lack of AIM-A1 in the crystals mean the interface between the nanobody and VWF can't be interrogated but the ND4 structure was instead utilised to provide other meaningful information. ND6 and ND4 crystal structures were overlaid to assess similarity between structures. The CDR loop 3 is the only region of visible

difference when the ND4 and 6 structures are overlaid with the ND4 nanobody with a larger looping region and small helical section along the face of the nanobody predicted to interact with AIM-A1. The nanobodies have 66% identity, with differences accounted for in CDR loops 1 and 3 as well as the terminal region (unresolved likely due to high flexibility) Figure 6.10.

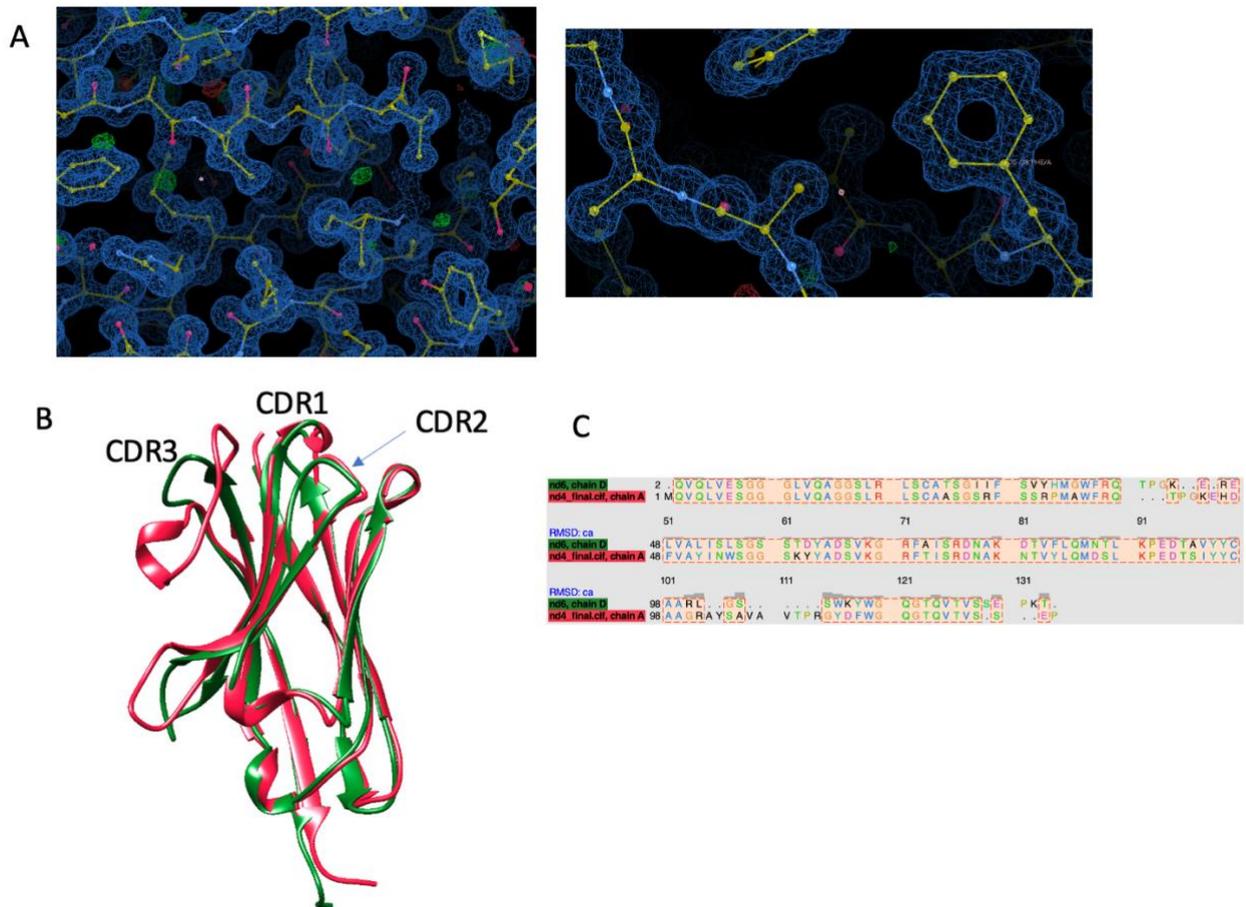


Figure 6.10: High resolution ND4 structure with comparison to ND6. (A) Electron density map at 1 Å resolution indicating the position of individual atoms of ND4 nanobody (B) Cartoon representation of ND6 (PDB code) in green and ND4 in red overlaid to visualise similarities between main chain and CDR loops. (C) Sequence alignment of ND4 and ND6 sequences, 66% identity seen.

6.2.8 AIM-A1 crystallisation trials

The VWF AIM-A1 construct was also set up for crystallisation trials alone. After many months, crystals could be seen in the JCSG+ screen (0.2M Zinc acetate dihydrate, 0.1 M Imidazole 8.0, 20 % w/v PEG 3000). This condition was

subsequently optimised to encompass 0.12- 0.22M Zinc acetate dihydrate, 8%-22% PEG-3000 and 0.1M Imidazole 8.0. Subsequently large crystal clusters known as ‘hedgehog crystals’ could be seen growing in multiple conditions (Figure 6.11A-B). These were not suitable for X-ray diffraction experiments but with further optimisation single crystals may be obtained.

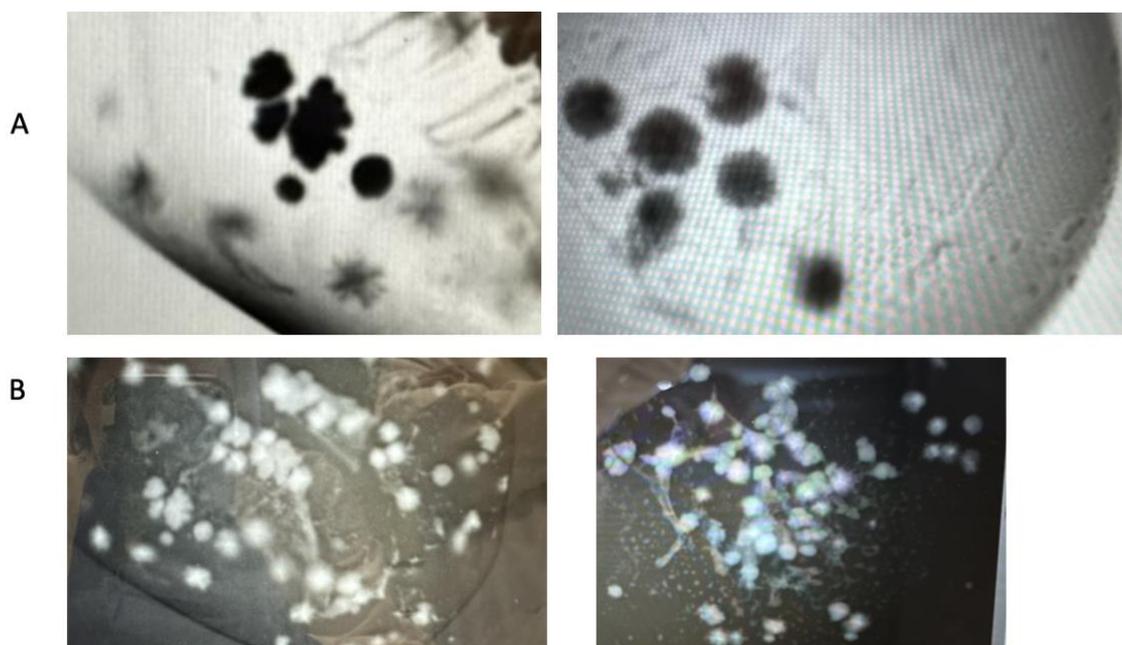


Figure 6.11: Images of crystals of VWF AIM-A1. Crystals of VWF AIM-A1 alone in 0.2M Zinc acetate dihydrate, 0.1 M Imidazole 8.0, 20% w/v PEG 3000, visualised utilising (A) light microscope and (B) UV-Visible Absorption microscope. Large crystal clusters can be seen formed of densely packed crystals unsuitable for X-ray diffraction experiments.

6.3 Discussion

Following the discovery of caplacizumab and its use in TTP, nanobody therapies are being further explored to target VWF AIM-A1. The make-up of the AIM-A1 was presented in schematic form in Figure 1.1, with the central A1 domain flanked by the flexible N and C terminal autoinhibitory modules. ND4 and ND6 were selected for their affinity specifically to the CAIM, with their binding affinity and subsequent crystal structures explored to understand their therapeutic utility to rival caplacizumab. The difference in the affinity of ND6 to VWF can likely be explained

by the different interactions with the CAIM; the interface between AIM and ND6 can now be interrogated whilst the ND4 complex still needs to be resolved. Although resolution of AIM-A1 ND6 is relatively low at 3.4Å, comparison to the high resolution ND4 structure and AIM-A1 (PDB: 1AUQ, 7A6O) structures, reveal large areas of similarity that provide confidence to the manual model building in areas where density is less defined (Figure 6.8 and 6.10). Improvements could be made to the data collection techniques to reduce radiation damage ensuring data is complete but also biologically relevant. The AIM-A1 ND6 complex structure provides crucial insight into both the interface of the complex, but perhaps more importantly, useful structural information on AIM-A1 and the potential mechanism of inhibition and activation.

An interesting newly resolved feature is that of the CAIM, specifically, the ‘PPPT’ residues from the PPPTLPP sequence (residues 1465-1471) are resolved more clearly than in previous structures, due to stabilisation of the CAIM by ND6. This site has been previously identified as the ristocetin binding site and although many studies make use of ristocetin binding, this area has not been visualised clearly (239). Sequences of repeating prolines have been explored as linkers, with utility for fusion protein design (240, 241). However, research into proline-rich sequences have also highlighted the importance of binding in these areas which can be more easily modulated than other sites. This is firstly due to their specific structure and also location on the edge of, or breaking up, domains (240). For example, 4-O-methyl-glucuronoyl methyltransferase 1 (A0A1D8EJG8) has repeats of PPPT. AlphaFold predicts these sequences form a flexible extending helical structure that loops round, providing support for the interesting conformation the CAIM and NAIM adopt at the termini of the AIM-A1 structure as well as being an area for modulation. This could be explored in greater detail along with other CAIM features, which perhaps allow for activation of A1 in response to specific changes in blood flow conditions.

There are still some unresolved residues at the far N and C-termini of AIM-A1 (1238-1257 and 1475-1493), it would be interesting to resolve these residues in future studies. The newly resolved residues in the NAIM and CAIM provide support for the unique nature of the AIMS in stabilisation of the A1 domain preventing activation until the shear force limit is surpassed. The A1-ND6 structure highlights

the co-operative nature of the AIMs, with the structure capturing a conformation suggestive of the proline-rich regions creating a U-shape effect. This brings the AIMs into proximity of each other suggesting interaction could occur at a separate interface away from the A1 domain itself (Figure 6.7). This feature has not been previously explored in detail likely due to previous structures not having stabilised AIMs, either due to lack of nanobody for stabilisation or missing O-glycosylation unable to give a complete indication of proximity.

Previous VWF A1 proteins used in crystallization were produced in *E.coli* and subsequently lacked this glycosylation visualized at residues T1468 and S1263. This may be an important structural feature related to the AIMs that has been previously overlooked (88). Current research into the glycosylation of AIMs of VWF suggest that the presence of the mono-sialylated core 1 glycans may affect activation of A1 via two potential mechanisms: through steric hindrance due to size of the glycan, and stabilising of the AIMs leading to reduced GPIIb α binding (88). In a recent paper the importance of the sialylation of these O-glycans was clear as removal led to increased activation of VWF AIM-A1 compared to WT (238). As suggested by Arce et al 2021, steric hindrance is provided by these now resolved CAIM residues, as well as the O-glycosylation to the GPIIb α LBD (68). It would be useful to obtain the crystal structure of AIM-A1 alone to compare this structure to the complex, specifically for details around glycosylation and the positioning of the flexible AIMs. This would provide both support to the mechanism of action of ND6, as well as providing a crystal structure of AIM-A1 in a more natural state with the glycosylation present which has not been achieved previously.

The AIM-A1 crystals obtained from crystallisation trials were too densely packed to provide any meaningful X-ray diffraction data but could provide a seeding stock to improve chances of further optimisation and crystal formation (242, 243).

Furthermore crystallisation conditions could be altered slightly to improve chances of seeding to gain single crystals based on microseed-matrix screening, as well as dilution of the seed stock to gain larger single crystals(244). The information obtained from this novel complex structure provides both a new avenue for exploration with nanobody treatments targeting the VWF A1 domain as well as elucidating new structural features of AIM-A1. Optimisations of the ND6 nanobody

to increase the binding affinity may be required to ensure this is a suitable therapeutic avenue to explore. Alternatively resolving the A1-ND4 structure may reveal an extended portion of the CAIM if sufficient stabilisation is achieved through tighter binding. If the extended CAIM and NAIM can be fully visualized in a future structure, it may provide critical evidence to confirm current theories that are forming into how the co-operative nature of the AIMS may occur and the role in inhibition of VWF activation.

7. Discussion and conclusions

TTP is an extremely heterogeneous disease, with clinical presentation, disease progression and severity differing widely between patients. Thus far it has been difficult to correlate TTP genotype to phenotype. Through analysis of ADAMTS13 variants, pathogenicity prediction is still suboptimal with over half the mutations analysed lacking a REVEL or SIFT score, indicating pathogenicity or aberrant protein structure respectively. Furthermore, when trying to ascertain associations between genotype and phenotype to enable better prediction of pathogenicity most analyses are limited by small sample size and inconsistent reporting of patient disease characteristics (131, 193). Mutations are present throughout ADAMTS13 domains, with a spread of neonatal and adult-onset disease, as well as structural changes seen in ADAMTS13. The results of the analysis of mutations here aligns with previous research with mutations around the active site commonly affecting activity of ADAMTS13. Furthermore, mutations in the CUB domains frequently affect protein stability and secretion due to the nature of the protein function associated with each domain (10, 179, 245). Besides this, the specific structural defect of the proteins seems to give a greater indication of disease severity with large structural changes causing a greater reduction in protein function or stability compared to the effect of the location of a mutation. It would be beneficial to conduct an up-to-date large scale analysis such as that in Alwan et al 2019, which suggested an association between pre-spacer mutations and earlier onset disease, to test this association while looking for others to improve predictions of disease progression as well as improving the initial diagnosis of cTTP to inform treatment decisions (38).

Achieving quick diagnosis of TTP is crucial for reduced time to initial treatment of patients. Issues with misdiagnosis of TTP are now less common with the wide availability of ADAMTS13 assays, of which multiple are now available, to distinguish TTP from other thrombosis disorders (36, 246). However, even with a confirmed ADAMTS13 diagnosis there still exists a lack of clear guidance towards best treatment course for TTP patients over the course of disease. The ISTH guidelines issue advice based on low certainty evidence in many scenarios (64). The analysis of these guidelines as well as patient mutations indicated whilst current

treatment options have dramatically improved mortality rates, there is still a need to develop further treatment options addressing issues of relapse and invasiveness of procedures to improve patient quality of life further. In order to have treatment options with higher certainty evidence, subsequently leading to universal agreement on treatment course, the understanding of the basic science behind TTP is needed. Investigating ADAMTS13 and the regulation of VWF and the dysregulation in disease will help inform the likely success of potential treatment avenues currently being explored.

To understand the pathophysiology behind TTP, understanding of the unique VWF and ADAMTS13 protein interaction is required. With protein-protein complex interactions between ADAMTS13 and VWF based on molecular modelling, the utility of AlphaFold here was explored for understanding latency and activation of ADAMTS13 and VWF but it is clear prediction software does not yet supersede experimental structures in these scenarios. The reliance of structure prediction software on similar protein structures in the PDB database limits all software from predicting, for example, the open structures of ADAMTS13 and the open linear structure of the VWF A2 domain (247). The use of these predicted structures of VWF constructs were however beneficial for the guiding of construct design. By analysing the predicted protein structure, regions of flexibility that may be determinantal to crystallisation of proteins were highlighted and subsequently removed or exploited to improve structural studies. Design of the 2GKG-internal-VWF structure allowed the TSP-1 repeat to be replaced with the stable 2GKG tag to aid in stability of the VWF A2 domain fragment whilst increasing points available for crystallisation contacts. Further utility for AlphaFold has been indicated for use in structure solution during molecular replacement both for X-ray crystallography and Cryogenic-electron microscopy experiments. The differences in AlphaFold models compared to the final structure following refinement, indicates the limitations of AlphaFold models alone currently however use as a complementary method can ease the process of structure determination (248).

Investigating the structure of other ADAMTS family proteins enabled design of a VWF construct to mimic the regulation of ADAMTS proteins by their propeptide. Whilst ADAMTS13 is unique in the redundancy of its propeptide, instead relying on

allosteric activation from VWF, the cysteine switch mechanism of metalloproteases can be exploited to aid in stable complex formation of ADAMTS13 (202, 249). The Y1605C VWF mutant shows increased affinity for ADAMTS13 compared to WT VWF constructs on preliminary ITC analysis likely due to this stronger hydrogen-bond with Zn^{2+} ion chelation of the cysteine. Addition of another cysteine (YM1605-6CC VWF mutant) lead to expression issues so caution should be taken when introducing cysteines to avoid protein destabilisation. Whilst the G1629E showed greater affinity to MDTCS (compared to WT) in *in vitro* studies previously, variable elution profiles were observed following size exclusion analysis (206). This mutation, in the context of the A2 domain fragment, may in fact decrease the likelihood of a stable complex formation with ADAMTS13 due to the unpredictable unstable nature of the linear peptide. It would still be beneficial to conduct preliminary binding analyses using SPR to confirm which constructs have utility going forward for stable ADAMTS13-VWF complex formation. The utility of the 2GKG tag on VWF constructs improved expression of the protein, which could be due to removal of flexible regions through utilising smaller sections of the VWF A2 domain. This still requires further investigation into complex formation with ADAMTS13 and comparing binding dynamics to other VWF constructs to select which constructs should be prioritised in future crystallisation trials.

The use of caplacizumab as a treatment for TTP created a much-needed alternative treatment to plasma exchange for cTTP patients. Since its introduction, issues have been highlighted with lack of efficacy in some patients, as well a potential for increased bleeding. This has led to a need for close monitoring of patients and required clinicians to have access to knowledge and equipment to offer caplacizumab as a viable treatment option, which may be difficult in low and middle-income countries paired with high unit costs (250). This subsequently led to a drive towards a nanobody treatment with both greater efficacy and safety to improve on the success of caplacizumab. The ND6 nanobody, selected through enrichment of AIM-A1 and then CAIM in yeast-phage display experiments, binds the CAIM with an affinity of 604nM (88, 234). Elucidation of the complex crystal structure of AIM-A1 and ND6 revealed both novel interaction information as well as a crucial insight into the role of the AIMS in VWF inhibition. The number and strength of interactions between AIM-A1 and ND6 varied compared to

caplacizumab, with a greater proportion of Van der Waals interactions observed with ND6 consistent with the reduced binding affinity compared to caplacizumab and AIM-A1 (88, 251). These observations from the crystal structure along with supplementary work conducted including HDX experiments as reported in Arce et al 2024 suggest that whilst ND6 and caplacizumab target alternative AIMS, the overall stabilisation of AIM-A1 is achieved by the same mechanism with subsequent inaccessibility of the GPIIb α binding site.

Insight into features of VWF regulation, in this case achieved artificially through use of a nanobody, helps understand the pathophysiology behind TTP but importantly, also the natural haemostasis and thrombosis processes. The interaction between the AIMS away from the A1 domain has been debated for multiple years. With this new structural evidence, the role in regulation of the AIM and exposure of the GPIIb α is clearer confirming the function of the AIMS rather than just linking the A1 domain to further regions of the VWF protein. It is also clear the use of a mammalian expression system is essential to truly understand the complex process occurring during activation of VWF, with O-glycosylation sites revealed for the first time on the AIM domains. The role of these O-glycans had been previously revealed through their desialylation leading to destabilisation of the AIMS and subsequent VWF activation, but until now the close proximity of the AIMS and their glycosylation at an interface away from the main A1 domain was not clear (238). Novel approaches are also needed to improve resolution of the A1 glycosylation, which is difficult due to their heterogeneity. Similar attempts at stabilising glycoproteins have been reported previously with a detailed explanation of issues reported in Kwong et al 1999 (218). Attempts to overcome missing glycosylation information across research areas followed a similar method to the addition of O-glycans on the A1 domain, and utilise modelling approaches based on previous research into structure of glycans and expected locations (252).

Increased understanding the methods of nanobody treatment, such as caplacizumab, in treating TTP by regulation of VWF is helping to progress nanobody therapeutics as a strong treatment option. Having multiple viable nanobody treatments can only be beneficial, helping to address issues of high costs, but also creating alternative treatment option for patients suffering relapse or with sub-optimal response to other

treatments. Progress is also being made in development of alternative treatments for TTP such as replacement therapy, targeting VWF regulation through restoring of ADAMTS13 function with a recombinant ADAMTS13 protein (253, 254) Thus far results from trials have provided promising results with good efficacy and safety outcomes (72). Long-term effects including safety and relapse rates have not been assessed, so continuing research into multiple treatment avenues should reveal information allowing informed decisions on best treatment options for TTP patients.

7.1 Study limitations and future work

The main limitation to these studies was the incomplete analysis of binding kinetics and complexing between ADAMTS13 and VWF as well as the lack of repeats in these experiments. Both issues were related to insufficient amount of ADAMTS13 protein generated due to unforeseen circumstances as well as time constraints. Both ADAMTS13 and VWF proteins contain regions of flexibility which led to difficulties in both expression and purification. The flexible linear region of VWF-73 constructs (specifically the 6xHis-SUMO VWF constructs) led to unexpected elution volumes when using size exclusion analysis and subsequently meant visualising of complexes with ADAMTS13 was not easily defined. The 2GKG constructs were designed to alleviate issues with the linear VWF A2 fragments as well as introducing a stable tag suitable for crystallisation, following experiments will reveal if this design is successful in addressing these issues. Alongside this, the ADAMTS13 construct (MDTCS E225Q) expression utilising S2 cells produced a lower yield than expected following induction and expression. Previous MDTCS E225Q protein expression was carried out in this expression system, with suitable levels of protein achieved for crystallisation trials so it is not clear why expression levels did not match those seen previously following the same method (97). It would be beneficial to repeat induction and expression optimisation experiments to confirm the optimal dosage of CuSO_4 and time for induction of protein expression. The analysis of expression would also be improved with the use of specific antibody for western blotting experiments. Both ADAMTS13 and VWF identification relied on the use of Anti-6xHis detection, so having specific antibodies would give more confidence with clearer detection of proteins.

The information from these expression and structural studies of ADAMTS13 and VWF informs construct design for generation of a stable complex between proteins. Whilst the Y1605C VWF mutant shows increased binding affinity to ADAMTS13 in preliminary analysis the 2GKG-constructs had better expression and purification with greater potential for crystallisation. Therefore, it could be worth combining these approaches in construct design if the binding analysis between ADAMTS13 and the 2GKG-VWF fragments (containing the scissile bond) is suboptimal. These constructs will reveal the structure of the active form of ADAMTS13, enabling visualisation of the local mechanism of latency. In order to visualise the global latency that regulates ADAMTS13 cleavage of VWF, the DTCS-CUB construct needs further investigation to design a construct capable of forming a stable Spacer-CUB interaction. Formation of a complex between VWF and DTCS will help elucidate this complex cycle of latency whilst also understanding how ADAMTS13 domains are altered by VWF binding.

To better understand VWF regulation and how autoinhibition is maintained by the A1 domain, the next step is further crystallisation of AIM-A1, both alone and in complex with an activating nanobody, to help visualise the full picture of activation. The role of the AIMS in this process will subsequently have sufficient evidence for inhibition of the A1 domain, disproving previous theories suggesting they are not of functional relevance and are merely structural linkers between domains. The research conducted here provides sound evidence towards the role of AIMS in inhibition of VWF due to specific interactions with this ND6 nanobody (alongside activating nanobodies utilised in complementary research studies) and AIMS, achieved through manipulation of the activation and inhibition of the A1 domain whilst only interacting with the AIMS (88).

7.2 Concluding remarks.

Sufficient progress has been made into construct design for VWF and ADAMTS13 constructs to understand the regulation of VWF by ADAMTS13. Whilst the exact interactions and activation states of ADAMTS13 and VWF were not yet elucidated, the predicted interaction information through docking as well as expression and purification of constructs provide a strong basis for experiments to obtain the complex crystal structure of ADAMTS13 and VWF required to visualise the

activation of these proteins. This information will be critical in increasing the understanding of how these proteins are affected in disease and inform disease management and treatment of TTP and perhaps wider haemostasis disorders. The VWF-ND6 crystal structure has provided evidence for the role of AIMS (and their glycosylation) in the inhibition of VWF activation, whilst also indicating the potential use of new nanobody treatments for TTP. This information should be useful in future investigation of aberrant VWF regulation and pathophysiological processes involving both VWF and ADAMTS13.

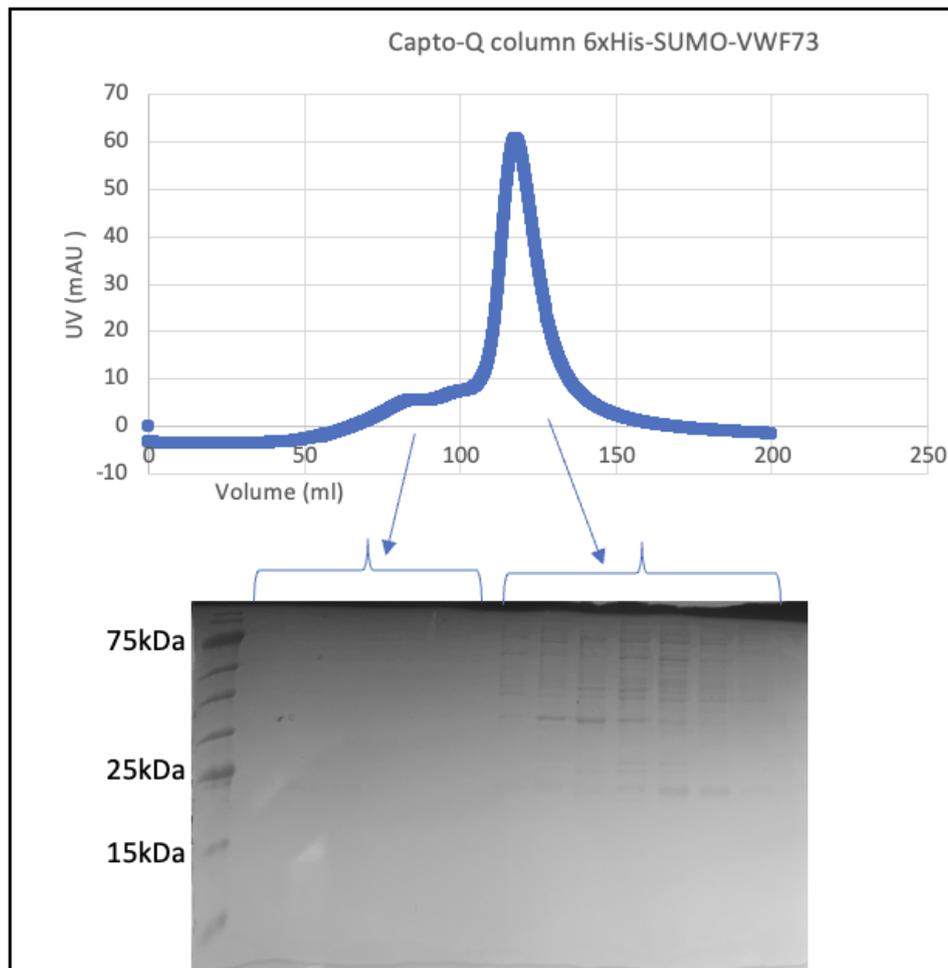
8. Supplementary data

Supplementary Table 1 Construct sequences. Sequences of ADAMTS13, VWF, and Nanobody construct sequences used in experiments.

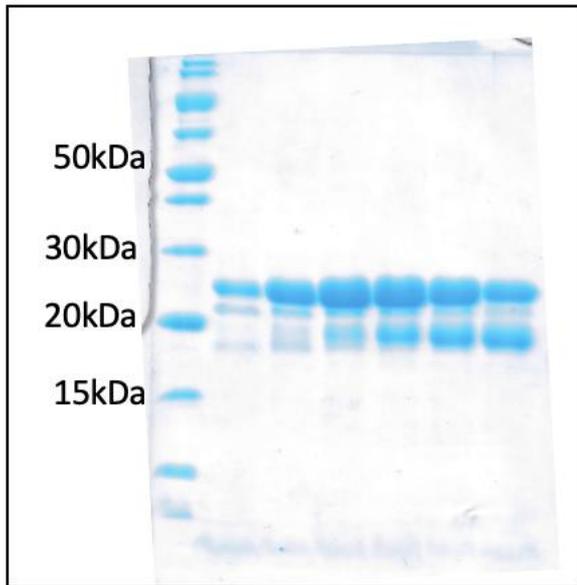
Construct name	Amino acid sequence
6xHis-VWF-73 WT	MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLK VSDGSSEIFFKIKKTTPLRRLMEAFKRQKGEMDSLRFYDGIQADQQT PEDLDMEDNDIIEAHREQIGGDREQAPNLVYMTGNPASDEIKRLPGDIQ VPIGVGPNANVQELERIGWPNAPILIQDFETLPREAPDLVLQR
6xHis-VWF-73 Y1605C	MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKV SDGSSEIFFKIKKTTPLRRLMEAFKRQKGEMDSLRFYDGIQADQQTPE DLDMEDNDIIEAHREQIGGDREQAPNLVCMVTGNPASDEIKRLPGDIQVV PIGVGPNANVQELERIGWPNAPILIQDFETLPREAPDLVLQR
6xHis-VWF-73 G1629E	MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKV SDGSSEIFFKIKKTTPLRRLMEAFKRQKGEMDSLRFYDGIQADQQTPE DLDMEDNDIIEAHREQIGGDREQAPNLVYMTGNPASDEIKRLPGDIQVV PIEVGPNANVQELERIGWPNAPILIQDFETLPREAPDLVLQR
6xHis-VWF-73 Y1605C+ G1629E	MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKV SDGSSEIFFKIKKTTPLRRLMEAFKRQKGEMDSLRFYDGIQADQQTPE DLDMEDNDIIEAHREQIGGDREQAPNLVCMVTGNPASDEIKRLPGDIQVV PIEVGPNANVQELERIGWPNAPILIQDFETLPREAPDLVLQR
6xHis-VWF-73 YM1605-6CC	MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKV SDGSSEIFFKIKKTTPLRRLMEAFKRQKGEMDSLRFYDGIQADQQTPE DLDMEDNDIIEAHREQIGGDREQAPNLVCCVTGNPASDEIKRLPGDIQVV PIGVGPNANVQELERIGWPNAPILIQDFETLPREAPDLVLQR
2GKG-internal-VWF73	MDREQAPNLVYMTGNPASDEIKRLPGDKKILIVESDTALSATLRSAL GRGFTVDETTDGGKSVEQIRDRPDLVLAVDLSAGQNGYLICGKLLK DDDLKNVPIVIGNPDGFAQHRKLLKAHADEYVAKPVDADQLVERAGAL IGFPIGWPNAPILIQDFETLPREAPDLVLQRGHHHHHHH
2GKG-MP-S-VWF	MKKILIVESDTALSATLRSALGRGFTVDETTDGGKSVEQIRDRPDLV LAVDLSAGQNGYLICGKLLKDDDLKNVPIVIGNPDGFAQHRKLLKAHAD EYVAKPVDADQLVERAGALIGFPRVYMTGNPASDEIKRLPGDGH HHH
2GKG-MP-L-VWF	MKKILIVESDTALSATLRSALGRGFTVDETTDGGKSVEQIRDRPDLV LAVDLSAGQNGYLICGKLLKDDDLKNVPIVIGNPDGFAQHRKLLKAHAD EYVAKPVDADQLVERAGALIGFPPNLVYMTGNPASDEIKRLPGDGH HHH
Cys-Spacer-2GKG-VWF	MKKILIVESDTALSATLRSALGRGFTVDETTDGGKSVEQIRDRPDLV VLAVDLSAGQNGYLICGKLLKDDDLKNVPIVIGNPDGFAQHRKLLKAH ADEYVAKPVDADQLVERAGALIGFPIGWPNAPILIQDFETLPREAPDLV LRGHHHHHHH
MDTCS E225Q	ILHLELLVAVGPDVFAQHQEDTERYVLTNLNIGAEALLRDP SLGAQFRVHLVKMVLTEPEGAPNITANLTSSLLSVC GWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTQLGGAC SPTWSCLITEDTGFDLGVITIAHQIGHSFGL EHDGAPGSGCGPSGHVMASDGAAPRAGLAWSP CSRRLQLLSLSSAGRARCWVDP PPRQPGSAGHPPDAQPGLYYSANEQCRVAFGPKAVAC TFAREHLDMCQALSCHTDPLDQSSCSRLLVPLLDGTECGVEK WCSKGRCRSLVELTPIAAVHGRWSSWGPRSPCSRSCGGGVVTRRRQC NNPRPAFGGRACV GADLQAEMCNTQACEKTQLEFMSQQCARTDGGQPLRSSPGGAS FYHWGAAVPHSQGDALCRHMCRAIGESFIMKRGDSFLDGT RCMPSPGPREDGTL SLCVSGSCRTFGCDGRMDSQQVWDRQCQVCGGDNSTC SPRKGSTAGRAREYVTF LTVTPNLTSVYIANHRPLFTHLAVRIGGRYV VAGKMSISPNNTYPSLLEDGRVEYRVALTEDRL PRLEEIRIWGPLQEDADIQVYRRYGE EYGNLTRPDITFTYFQPKP

DTCS-CUB	RSHHHHHHDMDGSIPGWKPLYGGADEQCKIAFGSVATACTFADSNVDI CEVLSCHVQPGDESSCARLLVPLLDGTECGINKWCSKGCSSLEELNPMA VHGQWSVWSPFSSCSRSCGGGVLRQRFCCNNPRPEPARWRCPVADSGP WGLSPALCLGQTCLQACSVTEQDFMAEQCAATNLKPLYLTGEAPSFYNW TSAVGFAGDMLCKHMCRAMGNEFMVSRKGSFIDGTRCEQDESERRGA FNLCVTGTCTRAFCDGQMDSKKIMDSCKVCGDNTTCTKVSYSYTEGKA KEYVTFLSLPYNTTSVHVNTNRPLFTHLAVKVKGEYAVAGKGKISQNV YPSVLEDSQIYKVFLTEDNLPSLEEIHVDGPTQEEIEIQVYRRYAKEYGNA TNPDITFSYFVPKENLTYVWIPQRGPCSVTCGEGERQLQYVCVAFDTEET QEENCRPVKPESRMEICDLRPCPPRWKVTAPGPCSSSCGLGLAVQLVTC VQIHQGKEMLLEERLCPVAEKPLTSVPCVIRMCSYEWFSFEWTECSTSCG NGIQTRQDFCLNPLTRKQVNPVFCRHFPAKIVVRGCSAGPCPEQAMGTV SHGAGLQVTPALPTTTTTAKEARYKDLPLPPSAVPVPEEAQETGAGV CGKFLNATGVINMTGVESDCTVAIGRPLGEEITLRVLESSLNCSAGEIV LFSGRMMWRTGCRNLPLSLINSRTNTLIVKQRVSLPGNGVILQYNSRTA TKKYYQDCDKQLFGPHGEIANPVQLPVQRQEVVCRTFINVAPRHRVAIR ALNIDLGNESNQTHFNILVRDVSTMKTMVFGKQOFFWQSTGSQAEIE FHENVKDYQTSFWAEYYAIEPKTR
ADAMTS13 human FL	MHQRHPRARCPPLCVAGILACGFLGCGWGPSHFQQSCLQALEPQAVSSY LSPGAPLKGRPPSPGFQRQRQRRAAGGILHELLVAVGPDVFAQHQE DTERYVLTNLNIGAEELLRDP SLGAQFRVHLVKMVILTEPEGAPNITANLT SSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELPDGNRQVRGVTQ LGGACPTWCLITEDTGFDLGVITIAHEIGHSFGLHGDGAPGSGCGPSGH VMASDGAAPRAGLAWSPCSRRLSLLSAGRARCVDPPRPQPGSAGH PPDAQPLYYSANEQCRVAFGPKAVACTFAREHLMCQALSCHTDPLD QSSCSRLLVPLLDGTECGVEKWCSKGRCSRSLVELTPIAAVHGRWSSWGP RSPCSRSCGGGVVTRRRQCNNPRPAFGGRACVGADLQAEMCNTQACEK TQLEFMSQQCARTDGGQLRSSPGGASFYHWGAAVPHSQGDALCRHMCR AIGESFIMKRGDSFLDGTTRCMPSPGPREDTLSLCVSGSCRTFGCDGRMDS QQVWDRQCQVCGGDNSTCSPRKGSTAGRAREYVFTLTVPNLTSVYIAN HRPLFTHLAVRIGGRYVAGKMSISPNTTYPSSLLEDGRVEYRVALTEDRL PRLEEIRIWGPLQEDADIQVYRRYGEEYGNLTRPDITFTYFQPKPRQAWV WAAVRGPCSVSCGAGLRWVNYSCLDQARKELVETVQCQGSQQPPAWP EACVLEPCPPYWAVGDFGPCSASCGGLRERPVRCVEAQGSLLKTLPPA RCRAGAQQPAVALETCNPQPCPARWEVSEPSCTAGGAGLALANETCV PGADGLEAPVTEGPGSVDEKLPAPEPCVGMSCPPGWGHLDATSAGEKAP SPWGSIRTGAQAAHVWTPAAGSCSVSCGRGLMELRFLCMDALRVPVQ EELCGLASKPGSRREVCQAVPCPARWQYKLAACSVSCGRGVVRRILYCA RAHGEDDGEEILLDTQCQGLPRPEPQEACSLPCPPRWKVMVSLGPCSASC GLGTARRSVACVQLDQGDVEVDEAACAALVRPEASVPCLIADCTYRW HVGWMECSVSCGDGIQRRRDTCLGPQAQAPVPADFCQHLPKPVTVRG CWAGPCVGGTTPSLVPHEEAAAPGRTTATPAGASLEWSQARGLLFSPAP QPRRLLPGQENSQSSACGRQHLEPTGTIDMRGPGQADCAVAIGRPLGE VTLRVLESSLNCSAGDMLLLWGRLTWRKMCRLKLLDMTFSSKTNTLVV RQRCGRPGGGVLLRYGSQLAPETFYRECDMQLFGPWGEIVSPSLSPATSN AGGCRLFINVAPHARIAIHALATNMGAGTEGANASYILIRDTHSLRTTAF HGQQVLYWESESSQAEMEFSEGFLKAQASLRGQYWTLQSWVPEMQDP QSWKGKEGT
VWF AIM- A1	GSQEPGGLVVPPTDAPVSPTTLYVEDISEPPLHDFYCSRLLDLVFLLDGSS RLSEAEFEVLKAFVVDMMERLRISQKWVRVAVVEYHDGSHAYIGLKDR KRPSLRRIASQVKYAGSQVASTSEVLKYTLFQIFSKIDRPEASRIALLM ASQEPQRMSRNFVRYVQGLKKKKVIVIPVGIGPHANLKQIRLIEKQAPEN KAFVLSVDELEQQRDEIVSYLCLDAPEAPPPTLPPDMAQVTVGPG

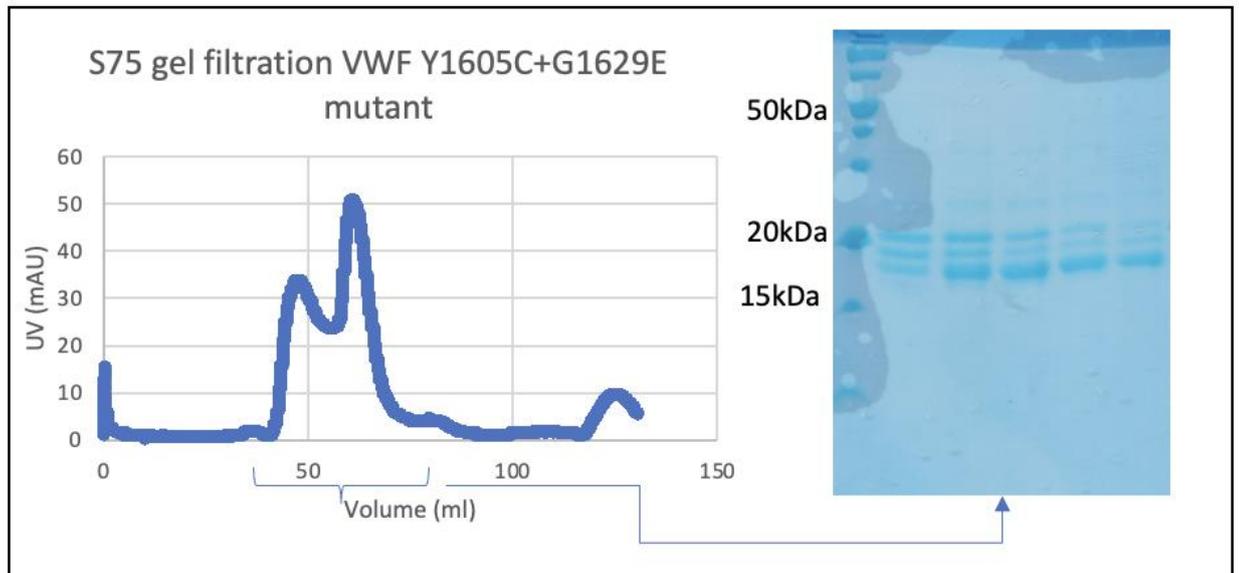
ND6	MQVQLVESGGGLVQAGGSLRLSCATSGIIFSVYHMGWFRQTPGKEREL VALISLSGSSTDYADSVKGRFAISRDNADKDTVFLQMNTLKPEDTAVYY CAARLGSSWKYWGQGTQVTVSSEPKTPKQPLEHHHHHH
ND4	MQVQLVESGGGLVQAGGSLRLSCAASGSRFSSRPMAWFRQTPGKEHDF VAYINWSSGSKYYADSVKGRFTISRDNADKNTVYLQMDSLKPEDTSIYY CAAGRAYSAAVAVTPRGYDFWGQGTQVTVSSHHHHHH



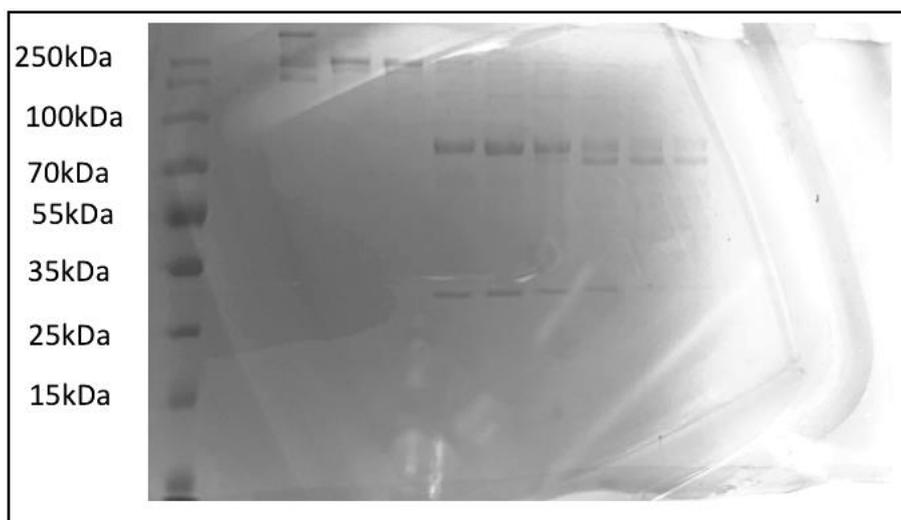
Supplementary figure 1: WT 6xHis-SUMO-VWF73 Capto-Q ion exchange trial. Analysis of WT VWF construct utilising Capto-Q ion exchange column did not result in increased yield of protein.



Supplementary figure 2: SDS-PAGE gel of 6xHis-SUMO-VWF73 WT after size exclusion chromatography at 0.8ml/min. Complete separation of bands has not been achieved at this speed.



Supplementary Figure 3: S75 elution profile and SDS-PAGE gel (containing the peak fractions) of Y1605C+G1629E mutant 6xHis-SUMO-VWF73 highlighting issues with expression and purification of a pure stable protein with multiple bands between 15-20kDa.



Supplementary figure 4: SDS-PAGE gel of fractions following DTCS-CUB S200 purification. Band at ~95kDa excised for Mass Spectrometry analysis.

References

1. Gale AJ. Current understanding of hemostasis. *Toxicol Pathol.* 2011;39(1):273-80.
2. Smith SA, Travers RJ, Morrissey JH. How it all starts: Initiation of the clotting cascade. *Crit Rev Biochem Mol Biol.* 2015;50(4):326-36.
3. Peyvandi F, Garagiola I, Baronciani L. Role of von Willebrand factor in the haemostasis. *Blood Transfusion.* 2011;9(SUPPL. 2):3-8.
4. Fowler WE, Fretto LJ, Hamilton KK, Erickson HP, McKee PA. Substructure of human von Willebrand factor. *Journal of Clinical Investigation.* 1985;76(4):1491-500.
5. Springer TA. Von Willebrand factor, Jedi knight of the bloodstream. *Blood.* 2014;124(9):1412-25.
6. Ulrichs H, Udvardy M, Lenting PJ, Pareyn I, Vandeputte N, Vanhoorelbeke K, et al. Shielding of the A1 domain by the D'D3 domains of von Willebrand factor modulates its interaction with platelet glycoprotein Ib-IX-V. *J Biol Chem.* 2006;281(8):4699-707.
7. Xu ER, von Bulow S, Chen PC, Lenting PJ, Kolsek K, Aponte-Santamaria C, et al. Structure and dynamics of the platelet integrin-binding C4 domain of von Willebrand factor. *Blood.* 2019;133(4):366-76.
8. Federici AB, Bader R, Pagani S, Colibretti ML, De Marco L, Mannucci PM. Binding of von Willebrand factor to glycoproteins Ib and IIb/IIIa complex: affinity is related to multimeric size. *British Journal of Haematology.* 1989;73(1):93-9.
9. Crawley JTB, De Groot R, Xiang Y, Luken BM, Lane DA. Unraveling the scissile bond: How ADAMTS13 recognizes and cleaves von Willebrand factor. *Blood* 2011. p. 3212-21.
10. Levy GG, Nichols WC, Lian EC, Foroud T, McClintick JN, McGee BM, et al. Mutations in a member of the ADAMTS gene family cause thrombotic thrombocytopenic purpura. *Nature: Nature;* 2001. p. 488-94.
11. Mead TJ, Apte SS. ADAMTS proteins in human disorders. *Matrix Biology: Elsevier B.V.;* 2018. p. 225-39.
12. Zhu J, Muia J, Gupta G, Westfield LA, Vanhoorelbeke K, Tolia NH, et al. Exploring the "minimal" structure of a functional ADAMTS13 by mutagenesis and small-angle X-ray scattering. *Blood.* 2019;133(17):1909-18.
13. Kelwick R, Desanlis I, Wheeler GN, Edwards DR. The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family. *Genome Biology.* 2015;16(1).
14. Zhou W, Inada M, Lee TP, Benten D, Lyubsky S, Bouhassira EE, et al. ADAMTS13 is expressed in hepatic stellate cells. *Lab Invest.* 2005;85(6):780-8.
15. Shelat SG, Ai J, Zheng XL. Molecular Biology of ADAMTS13 and Diagnostic Utility of ADAMTS13 Proteolytic Activity and Inhibitor Assays. *Seminars in thrombosis and hemostasis.* 2005;31(6):659-.
16. Turner N, Nolasco L, Tao Z, Dong JF, Moake J. Human endothelial cells synthesize and release ADAMTS-13. *J Thromb Haemost.* 2006;4(6):1396-404.
17. Suzuki M, Murata M, Matsubara Y, Uchida T, Ishihara H, Shibano T, et al. Detection of von Willebrand factor-cleaving protease (ADAMTS-13) in human platelets. *Biochem Biophys Res Commun.* 2004;313(1):212-6.

18. Ricketts LM, Dlugosz M, Luther KB, Haltiwanger RS, Majerus EM. O-fucosylation is required for ADAMTS13 secretion. *J Biol Chem.* 2007;282(23):17014-23.
19. Zhou W, Tsai HM. N-Glycans of ADAMTS13 modulate its secretion and von Willebrand factor cleaving activity. *Blood.* 2009;113(4):929-35.
20. Ling J, Su J, Ma Z, Ruan C. The WXXW motif in the TSR1 of ADAMTS13 is important for its secretion and proteolytic activity. *Thrombosis Research.* 2013;131(6):529-34.
21. Verbij FC, Stokhuijzen E, Kaijen PH, van Alphen F, Meijer AB, Voorberg J. Identification of glycans on plasma-derived ADAMTS13. *Blood.* 2016;128(21):e51-e8.
22. Furlan M, Robles R, Morselli B, Sandoz P, Lammle B. Recovery and half-life of von Willebrand factor-cleaving protease after plasma therapy in patients with thrombotic thrombocytopenic purpura. *Thromb Haemost.* 1999;81(1):8-13.
23. Rieger M, Ferrari S, Kremer Hovinga JA, Konetschny C, Herzog A, Koller L, et al. Relation between ADAMTS13 activity and ADAMTS13 antigen levels in healthy donors and patients with thrombotic microangiopathies (TMA). *Thromb Haemost.* 2006;95(2):212-20.
24. Soejima K, Nakamura H, Hirashima M, Morikawa W, Nozaki C, Nakagaki T. Analysis on the molecular species and concentration of circulating ADAMTS13 in Blood. *J Biochem.* 2006;139(1):147-54.
25. Longpre JM, McCulloch DR, Koo BH, Alexander JP, Apte SS, Leduc R. Characterization of proADAMTS5 processing by proprotein convertases. *Int J Biochem Cell Biol.* 2009;41(5):1116-26.
26. Majerus EM, Zheng X, Tuley EA, Sadler JE. Cleavage of the ADAMTS13 Propeptide Is Not Required for Protease Activity. *Journal of Biological Chemistry.* 2003;278(47):46643-8.
27. Jones GC, Riley GP. ADAMTS proteinases: a multi-domain, multi-functional family with roles in extracellular matrix turnover and arthritis. *Arthritis Res Ther.* 2005;7(4):160-9.
28. Colige A, Sieron AL, Li SW, Schwarze U, Petty E, Wertelecki W, et al. Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene. *Am J Hum Genet.* 1999;65(2):308-17.
29. Brouillard P, Dupont L, Helaers R, Coulie R, Tiller GE, Peeden J, et al. Loss of ADAMTS3 activity causes Hennekam lymphangiectasia-lymphedema syndrome 3. *Hum Mol Genet.* 2017;26(21):4095-104.
30. Bauer RC, Tohyama J, Cui J, Cheng L, Yang J, Zhang X, et al. Knockout of Adamts7, a novel coronary artery disease locus in humans, reduces atherosclerosis in mice. *Circulation.* 2015;131(13):1202-13.
31. Larkin J, Lohr TA, Elefante L, Shearin J, Matico R, Su JL, et al. Translational development of an ADAMTS-5 antibody for osteoarthritis disease modification. *Osteoarthritis Cartilage.* 2015;23(8):1254-66.
32. Moschcowitz E, editor Hyaline thrombosis of the terminal arterioles and capillaries: a hitherto undescribed disease. *Proc NY Pathol Soc;* 1924.

33. Upshaw JD, Jr. Congenital deficiency of a factor in normal plasma that reverses microangiopathic hemolysis and thrombocytopenia. *N Engl J Med.* 1978;298(24):1350-2.
34. Schulman I, Pierce M, Lukens A, Currimbhoy Z. Studies on thrombopoiesis. I. A factor in normal human plasma required for platelet production; chronic thrombocytopenia due to its deficiency. *Blood.* 1960;16:943-57.
35. Stanley M, Killeen RB, Michalski JM. Thrombotic Thrombocytopenic Purpura. *StatPearls.* Treasure Island (FL)2023.
36. Chiasakul T, Cuker A. Clinical and laboratory diagnosis of TTP: an integrated approach. *Hematology Am Soc Hematol Educ Program.* 2018;2018(1):530-8.
37. Scully M, Thomas M, Underwood M, Watson H, Langley K, Camilleri RS, et al. Thrombotic thrombocytopenic purpura and pregnancy: presentation, management, and subsequent pregnancy outcomes. *Blood.* 2014;124(2):211-9.
38. Alwan F, Vendramin C, Liesner R, Clark A, Lester W, Dutt T, et al. Characterization and treatment of congenital thrombotic thrombocytopenic purpura. *Blood.* 2019;133(15):1644-51.
39. Patel J, Patel P, Ahmed Z. An improbable and unusual case of thrombotic thrombocytopenia purpura. *J Community Hosp Intern Med Perspect.* 2016;6(4):32258.
40. Bell WR, Braine HG, Ness PM, Kickler TS. Improved survival in thrombotic thrombocytopenic purpura-hemolytic uremic syndrome. Clinical experience in 108 patients. *N Engl J Med.* 1991;325(6):398-403.
41. Tsai HM, Lian EC. Antibodies to von Willebrand factor-cleaving protease in acute thrombotic thrombocytopenic purpura. *N Engl J Med.* 1998;339(22):1585-94.
42. Underwood MI, Alwan F, Thomas MR, Scully MA, Crawley JTB. Autoantibodies enhance ADAMTS-13 clearance in patients with immune thrombotic thrombocytopenic purpura. *J Thromb Haemost.* 2023;21(6):1544-52.
43. Matsumoto M, Miyakawa Y, Kokame K, Ueda Y, Wada H, Higasa S, et al. Diagnostic and treatment guidelines for thrombotic thrombocytopenic purpura (TTP) in Japan 2023. *Int J Hematol.* 2023;118(5):529-46.
44. Mariotte E, Azoulay E, Galicier L, Rondeau E, Zouiti F, Boisseau P, et al. Epidemiology and pathophysiology of adulthood-onset thrombotic microangiopathy with severe ADAMTS13 deficiency (thrombotic thrombocytopenic purpura): A cross-sectional analysis of the French national registry for thrombotic microangiopathy. *The Lancet Haematology.* 2016;3(5):e237-e45.
45. Oka S, Nohgawa M. EB virus reactivation triggers thrombotic thrombocytopenic purpura in a healthy adult. *Leuk Res Rep.* 2017;8:1-3.
46. Kosugi N, Tsurutani Y, Isonishi A, Hori Y, Matsumoto M, Fujimura Y. Influenza A infection triggers thrombotic thrombocytopenic purpura by producing the anti-ADAMTS13 IgG inhibitor. *Intern Med.* 2010;49(7):689-93.
47. Motto DG, Chauhan AK, Zhu G, Homeister J, Lamb CB, Desch KC, et al. Shigatoxin triggers thrombotic thrombocytopenic purpura in genetically susceptible ADAMTS13-deficient mice. *J Clin Invest.* 2005;115(10):2752-61.
48. Holdrinet R, Pauw Bd, Haanen C. Hormonal dependent thrombotic thrombocytopenic purpura (TTP). *Scandinavian journal of haematology.* 1983;30(3):250-6.

49. Liang R, Wong RW, Cheng IK. Thrombotic thrombocytopenic purpura and 17 beta-estradiol transdermal skin patch. *Am J Hematol.* 1996;52(4):334-5.
50. Azarm T, Sohrabi A, Mohajer H, Azarm A. Thrombotic Thrombocytopenic Purpura associated with Clopidogrel: a case report and review of the literature. *J Res Med Sci.* 2011;16(3):353-7.
51. Scully M, Yarranton H, Liesner R, Cavenagh J, Hunt B, Benjamin S, et al. Regional UK TTP Registry: correlation with laboratory ADAMTS 13 analysis and clinical features. *British Journal of Haematology.* 2008;142(5):819-26.
52. Blombery P, Scully M. Management of thrombotic thrombocytopenic purpura: current perspectives. *J Blood Med.* 2014;5:15-23.
53. Bendapudi PK, Hurwitz S, Fry A, Marques MB, Waldo SW, Li A, et al. Derivation and external validation of the PLASMIC score for rapid assessment of adults with thrombotic microangiopathies: a cohort study. *Lancet Haematol.* 2017;4(4):e157-e64.
54. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res.* 2001;11(5):863-74.
55. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248-9.
56. Flanagan SE, Patch AM, Ellard S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet Test Mol Biomarkers.* 2010;14(4):533-7.
57. Huang B, Kong L, Wang C, Ju F, Zhang Q, Zhu J, et al. Protein Structure Prediction: Challenges, Advances, and the Shift of Research Paradigms. *Genomics Proteomics Bioinformatics.* 2023;21(5):913-25.
58. Ljungdahl A, Kohani S, Page NF, Wells ES, Wigdor EM, Dong S, et al. AlphaMissense is better correlated with functional assays of missense impact than earlier prediction algorithms. *bioRxiv.* 2023.
59. Rock GA, Shumak KH, Buskard NA, Blanchette VS, Kelton JG, Nair RC, et al. Comparison of plasma exchange with plasma infusion in the treatment of thrombotic thrombocytopenic purpura. Canadian Apheresis Study Group. *N Engl J Med.* 1991;325(6):393-7.
60. Hovinga JAK, Vesely SK, Terrell DR, Lammle B, George JN. Survival and relapse in patients with thrombotic thrombocytopenic purpura. 2010.
61. Shumak KH, Rock GA, Nair RC. Late relapses in patients successfully treated for thrombotic thrombocytopenic purpura. Canadian Apheresis Group. *Ann Intern Med.* 1995;122(8):569-72.
62. Vesely SK, George JN, Lammle B, Studt JD, Alberio L, El-Harake MA, et al. ADAMTS13 activity in thrombotic thrombocytopenic purpura-hemolytic uremic syndrome: relation to presenting features and clinical outcomes in a prospective cohort of 142 patients. *Blood.* 2003;102(1):60-8.
63. Sarode R, Bandarenko N, Brecher ME, Kiss JE, Marques MB, Szczepiorkowski ZM, et al. Thrombotic thrombocytopenic purpura: 2012 American Society for Apheresis (ASFA) consensus conference on classification, diagnosis, management, and future research. *J Clin Apher.* 2014;29(3):148-67.

64. Zheng XL, Vesely SK, Cataland SR, Coppo P, Geldziler B, Iorio A, et al. ISTH guidelines for treatment of thrombotic thrombocytopenic purpura. *Journal of Thrombosis and Haemostasis*. 2020;18(10):2496-502.
65. Sanofi. Sanofi: Cablivi(TM) (caplacizumab) approved in Europe for adults with acquired thrombotic thrombocytopenic purpura (aTTP) 2018 [Available from: <https://www.sanofi.com/en/media-room/press-releases/2018/2018-09-03-05-00-00-1564330>].
66. Sanofi. Sanofi: FDA approves Cablivi® (caplacizumab-yhdp), the first Nanobody®-based medicine, for adults with acquired thrombotic thrombocytopenic purpura (aTTP) 2019 [Available from: <https://www.sanofi.com/en/media-room/press-releases/2019/2019-02-06-16-43-21-1711610>].
67. Care DoHaA. Cablivi. 2020.
68. Arce NA, Cao W, Brown AK, Legan ER, Wilson MS, Xu ER, et al. Activation of von Willebrand factor via mechanical unfolding of its discontinuous autoinhibitory module. *Nat Commun*. 2021;12(1):2360.
69. Callewaert F, Roodt J, Ulrichs H, Stohr T, van Rensburg WJ, Lamprecht S, et al. Evaluation of efficacy and safety of the anti-VWF Nanobody ALX-0681 in a preclinical baboon model of acquired thrombotic thrombocytopenic purpura. *Blood*. 2012;120(17):3603-10.
70. Boothby A, Mazepa M. Caplacizumab for congenital thrombotic thrombocytopenic purpura. *Am J Hematol*. 2022;97(11):E420-E1.
71. Jin SY, Xiao J, Bao J, Zhou S, Wright JF, Zheng XL. AAV-mediated expression of an ADAMTS13 variant prevents shigatoxin-induced thrombotic thrombocytopenic purpura. *Blood*. 2013;121(19):3825-9.
72. ADZYNMA [package insert]. U.S. Food and Drug Administration website. (revised 11/2023).
73. Nov 2023 Summary Basis for Regulatory Action ADZYNMA. In: Administration) FFaD, editor. [Internet]2023.
74. A Study of BAX 930 in Children, Teenagers, and Adults Born With Thrombotic Thrombocytopenic Purpura (TTP) [Internet]. 2024. Available from: <https://clinicaltrials.gov/study/NCT03393975?term=NCT03393975&rank=1>.
75. A Study of TAK-755 in Participants With Congenital Thrombotic Thrombocytopenic Purpura [Internet]. 2024. Available from: <https://clinicaltrials.gov/study/NCT04683003>.
76. Goodeve AC. VWF sequence variants: a data goldmine. *Blood*. 2013;122(4):471-3.
77. Nowak AA, Canis K, Riddell A, Laffan MA, McKinnon TA. O-linked glycosylation of von Willebrand factor modulates the interaction with platelet receptor glycoprotein Ib under static and shear stress conditions. *Blood*. 2012;120(1):214-22.
78. Canis K, McKinnon TA, Nowak A, Haslam SM, Panico M, Morris HR, et al. Mapping the N-glycome of human von Willebrand factor. *Biochem J*. 2012;447(2):217-28.
79. McKinnon TA, Chion AC, Millington AJ, Lane DA, Laffan MA. N-linked glycosylation of VWF modulates its interaction with ADAMTS13. *Blood*. 2008;111(6):3042-9.

80. Gill JC, Endres-Brooks J, Bauer PJ, Marks WJ, Jr., Montgomery RR. The effect of ABO blood group on the diagnosis of von Willebrand disease. *Blood*. 1987;69(6):1691-5.
81. Wagner DD, Olmsted JB, Marder VJ. Immunolocalization of von Willebrand protein in Weibel-Palade bodies of human endothelial cells. *J Cell Biol*. 1982;95(1):355-60.
82. Anderson JR, Li J, Springer TA, Brown A. Structures of VWF tubules before and after concatemerization reveal a mechanism of disulfide bond exchange. *Blood*. 2022;140(12):1419-30.
83. Haberichter SL. von Willebrand factor propeptide: biology and clinical utility. *Blood*. 2015;126(15):1753-61.
84. Ruggeri ZM, Orje JN, Habermann R, Federici AB, Reininger AJ. Activation-independent platelet adhesion and aggregation under elevated shear stress. *Blood*. 2006;108(6):1903-10.
85. Schneider SW, Nuschele S, Wixforth A, Gorzelanny C, Alexander-Katz A, Netz RR, et al. Shear-induced unfolding triggers adhesion of von Willebrand factor fibers. *Proc Natl Acad Sci U S A*. 2007;104(19):7899-903.
86. Dumas JJ, Kumar R, McDonagh T, Sullivan F, Stahl ML, Somers WS, et al. Crystal structure of the wild-type von Willebrand factor A1-glycoprotein Ibalpha complex reveals conformation differences with a complex bearing von Willebrand disease mutations. *J Biol Chem*. 2004;279(22):23327-34.
87. Bonazza K, Iacob RE, Hudson NE, Li J, Lu C, Engen JR, et al. Von Willebrand factor A1 domain stability and affinity for GPIbalpha are differentially regulated by its O-glycosylated N- and C-linker. *Elife*. 2022;11.
88. Arce NA, Markham-Lee Z, Liang Q, Najmudin S, Legan ER, Dean G, et al. Conformational activation and inhibition of von Willebrand factor by targeting its autoinhibitory module. *Blood*. 2024.
89. Wiita AP, Ainaravapu SR, Huang HH, Fernandez JM. Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. *Proc Natl Acad Sci U S A*. 2006;103(19):7222-7.
90. Dong JF, Berndt MC, Schade A, McIntire LV, Andrews RK, Lopez JA. 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*. 2001;97(1):162-8.
91. Chen J, Ling M, Fu X, Lopez JA, Chung DW. Simultaneous exposure of sites in von Willebrand factor for glycoprotein Ib binding and ADAMTS13 cleavage: studies with ristocetin. *Arterioscler Thromb Vasc Biol*. 2012;32(11):2625-30.
92. Ajzenberg N, Ribba AS, Rastegar-Lari G, Meyer D, Baruch D. Effect of recombinant von Willebrand factor reproducing type 2B or type 2M mutations on shear-induced platelet aggregation. *Blood*. 2000;95(12):3796-803.
93. Azuma H, Sugimoto M, Ruggeri ZM, Ware J. A role for von Willebrand factor proline residues 702-704 in ristocetin-mediated binding to platelet glycoprotein Ib. *Thromb Haemost*. 1993;69(2):192-6.
94. Nishio K, Anderson PJ, Zheng XL, Sadler JE. Binding of platelet glycoprotein Ibalpha to von Willebrand factor domain A1 stimulates the cleavage of the adjacent domain A2 by ADAMTS13. *Proc Natl Acad Sci U S A*. 2004;101(29):10578-83.

95. De Groot R, Lane DA, Crawley JTB. The ADAMTS13 metalloprotease domain: Roles of subsites in enzyme activity and specificity. *Blood*. 2010;116(16):3064-72.
96. Xiang Y, De Groot R, Crawley JTB, Lane DA. Mechanism of von Willebrand factor scissile bond cleavage by a disintegrin and metalloproteinase with a thrombospondin type 1 motif, member 13 (ADAMTS13). *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(28):11602-7.
97. Petri A, Kim HJ, Xu Y, de Groot R, Li C, Vandenbulcke A, et al. Crystal structure and substrate-induced activation of ADAMTS13. *Nature Communications*. 2019;10(1).
98. Zhang Q, Zhou Y-F, Zhang C-Z, Zhang X, Lu C, Springer TA. Structural specializations of A2, a force-sensing domain in the ultralarge vascular protein von Willebrand factor. 2009.
99. Luken BM, Winn LY, Emsley J, Lane DA, Crawley JT. The importance of vicinal cysteines, C1669 and C1670, for von Willebrand factor A2 domain function. *Blood*. 2010;115(23):4910-3.
100. Zhou M, Dong X, Baldauf C, Chen H, Zhou Y, Springer TA, et al. A novel calcium-binding site of von Willebrand factor A2 domain regulates its cleavage by ADAMTS13. *Blood*. 2011;117(17):4623-31.
101. Goodeve AC. The genetic basis of von Willebrand disease. *Blood Rev*. 2010;24(3):123-34.
102. Fernández KS, de Alarcón PA. Von Willebrand Disease: Range of the Disease, and Management. *Current Pediatrics Reports*. 2014;2(1):60-70.
103. Atiq F, Wuijster E, de Maat MPM, Kruij M, Cnossen MH, Leebeek FWG. Criteria for low von Willebrand factor diagnosis and risk score to predict future bleeding. *J Thromb Haemost*. 2021;19(3):719-31.
104. Laffan MA, Lester W, O'Donnell JS, Will A, Tait RC, Goodeve A, et al. The diagnosis and management of von Willebrand disease: a United Kingdom Haemophilia Centre Doctors Organization guideline approved by the British Committee for Standards in Haematology. *Br J Haematol*. 2014;167(4):453-65.
105. James PD, Lillicrap D. The molecular characterization of von Willebrand disease: good in parts. *Br J Haematol*. 2013;161(2):166-76.
106. Lynch CJ, Cawte AD, Millar CM, Rueda D, Lane DA. A common mechanism by which type 2A von Willebrand disease mutations enhance ADAMTS13 proteolysis revealed with a von Willebrand factor A2 domain FRET construct. *PLoS ONE*. 2017;12(11).
107. Dong J, Zhao X, Shi S, Ma Z, Liu M, Wu Q, et al. Identification and functional analysis of a novel von Willebrand factor mutation in a family with type 2A von Willebrand disease. *PLoS ONE*. 2012;7(3).
108. Legan ER, Liu Y, Arce NA, Parker ET, Lollar P, Zhang XF, et al. Type 2B von Willebrand disease mutations differentially perturb autoinhibition of the A1 domain. *Blood*. 2023;141(10):1221-32.
109. Akiyama M, Takeda S, Kokame K, Takagi J, Miyata T. Crystal structures of the noncatalytic domains of ADAMTS13 reveal multiple discontinuous exosites for von Willebrand factor. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(46):19274-9.

110. Kim HJ, Xu Y, Petri A, Vanhoorelbeke K, B Crawley JT, Emsley J. Crystal structure of ADAMTS13 CUB domains reveals their role in global latency. 2021.
111. Ercig B, Wichapong K, Reutelingsperger CPM, Vanhoorelbeke K, Voorberg J, Nicolaes GAF. Insights into 3D Structure of ADAMTS13: A Stepping Stone towards Novel Therapeutic Treatment of Thrombotic Thrombocytopenic Purpura. *Thrombosis and Haemostasis*: Georg Thieme Verlag; 2018. p. 28-41.
112. Tallant C, Garcia-Castellanos R, Baumann U, Gomis-Ruth FX. On the relevance of the Met-turn methionine in metzincins. *J Biol Chem*. 2010;285(18):13951-7.
113. Barrett AJ, Woessner JF, Rawlings ND. *Handbook of Proteolytic Enzymes*, Volume 1: Elsevier; 2012.
114. De Groot R, Bardhan A, Ramroop N, Lane DA, Crawley JTB. Essential role of the disintegrin-like domain in ADAMTS13 function. *Blood*. 2009;113(22):5609-16.
115. Gao W, Anderson PJ, Majerus EM, Tuley EA, Sadler JE. Exosite interactions contribute to tension-induced cleavage of von Willebrand factor by the antithrombotic ADAMTS13 metalloprotease. 2006.
116. De Groot R, Lane DA, Crawley JTB. The role of the ADAMTS13 cysteine-rich domain in VWF binding and proteolysis. *Blood*. 2015;125(12):1968-75.
117. Soejima K, Matsumoto M, Kokame K, Yagi H, Ishizashi H, Maeda H, et al. ADAMTS-13 cysteine-rich/spacer domains are functionally essential for von Willebrand factor cleavage. *Blood*. 2003;102(9):3232-7.
118. Kokame K, Matsumoto M, Soejima K, Yagi H, Ishizashi H, Funato M, et al. Mutations and common polymorphisms in ADAMTS13 gene responsible for von Willebrand factor-cleaving protease activity. *Proceedings of the National Academy of Sciences of the United States of America*. 2002;99(18):11902-7.
119. Veyradier A, Lavergne JM, Ribba AS, Obert B, Loirat C, Meyer D, et al. Ten candidate ADAMTS13 mutations in six French families with congenital thrombotic thrombocytopenic purpura (Upshaw-Schulman syndrome). *Journal of Thrombosis and Haemostasis*. 2004;2(3):424-9.
120. Zheng X, Nishio K, Majerus EM, Sadler JE. Cleavage of von Willebrand Factor Requires the Spacer Domain of the Metalloprotease ADAMTS13*. *Journal of Biological Chemistry*. 2003;278:30136-41.
121. Graça NAG, Ercig B, Pereira LCV, Kangro K, Kaijen P, Nicolaes GAF, et al. Modifying ADAMTS13 to modulate binding of pathogenic autoantibodies of patients with acquired thrombotic thrombocytopenic purpura. *Haematologica*. 2020;105(11):2619-30.
122. Jian C, Xiao J, Gong L, Skipwith CG, Jin SY, Kwaan HC, et al. Gain-of-function ADAMTS13 variants that are resistant to autoantibodies against ADAMTS13 in patients with acquired thrombotic thrombocytopenic purpura. *Blood*. 2012;119(16):3836-43.
123. Tan K, Duquette M, Liu JH, Dong Y, Zhang R, Joachimiak A, et al. Crystal structure of the TSP-1 type 1 repeats: A novel layered fold and its biological implication. *Journal of Cell Biology*. 2002;159(2):373-82.
124. South K, Luken BM, Crawley JTB, Phillips R, Thomas M, Collins RF, et al. Conformational activation of ADAMTS13. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(52):18578-83.

125. Muia J, Zhu J, Gupta G, Haberichter SL, Friedman KD, Feys HB, et al. Allosteric activation of ADAMTS13 by von Willebrand factor. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(52):18584-9.
126. Deforche L, Roose E, Vandenbulcke A, Vandeputte N, Feys HB, Springer TA, et al. Linker regions and flexibility around the metalloprotease domain account for conformational activation of ADAMTS-13. *Journal of Thrombosis and Haemostasis*. 2015;13(11):2063-75.
127. Wang LW, Leonhard-Melief C, Haltiwanger RS, Apte SS. Post-translational modification of thrombospondin type-1 repeats in ADAMTS-like 1/punctin-1 by C-mannosylation of tryptophan. *J Biol Chem*. 2009;284(44):30004-15.
128. Zander CB, Cao W, Zheng XL. ADAMTS13 and von Willebrand factor interactions. *Current Opinion in Hematology*. 2015;22(5):452-9.
129. Muia J, Zhu J, Greco SC, Vanhoorelbeke K, Gupta G, Westfield LA, et al. Phylogenetic and functional analysis of ADAMTS13 identifies highly conserved domains essential for allosteric regulation. *Blood*. 2019;133(17):1899-908.
130. Licht C, Stapenhorst L, Simon T, Budde U, Schneppenheim R, Hoppe B. Two novel ADAMTS13 gene mutations in thrombotic thrombocytopenic purpura/hemolytic-uremic syndrome (TTP/HUS). *Kidney International*. 2004;66(3):955-8.
131. Lotta LA, Garagiola I, Palla R, Cairo A, Peyvandi F. ADAMTS13 mutations and polymorphisms in congenital thrombotic thrombocytopenic purpura. *Human Mutation*. 2010;31(1):11-9.
132. Banno F, Chauhan AK, Kokame K, Yang J, Miyata S, Wagner DD, et al. The distal carboxyl-terminal domains of ADAMTS13 are required for regulation of in vivo thrombus formation. *Blood*. 2009;113(21):5323-9.
133. Feys HB, Anderson PJ, Vanhoorelbeke K, Majerus EM, Sadler JE. Multi-step binding of ADAMTS-13 to von Willebrand factor. *Journal of Thrombosis and Haemostasis*. 2009;7(12):2088-95.
134. Halkidis K, Meng C, Liu S, Mayne L, Siegel DL, Zheng XL. Mechanisms of inhibition of human monoclonal antibodies in immune thrombotic thrombocytopenic purpura. *Blood*. 2023;141(24):2993-3005.
135. Taylor A, Vendramin C, Oosterholt S, Della Pasqua O, Scully M. Pharmacokinetics of plasma infusion in congenital thrombotic thrombocytopenic purpura. *Journal of Thrombosis and Haemostasis*. 2019;17(1):88-98.
136. Zhu L, Chen X, Abola EE, Jing L, Liu W. Serial Crystallography for Structure-Based Drug Discovery. *Trends Pharmacol Sci*. 2020;41(11):830-9.
137. Carugo O, Djinovic-Carugo K. Structural biology: A golden era. *PLoS Biol*. 2023;21(6):e3002187.
138. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-9.
139. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. 2021.
140. Yang Z, Zeng X, Zhao Y, Chen R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduction and Targeted Therapy*. 2023;8(1).

141. Terwilliger TC, Liebschner D, Croll TI, Williams CJ, McCoy AJ, Poon BK, et al. AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nat Methods*. 2023.
142. Singh K, Madarati H, Sohrabipour S, Sparring T, Teney C, Kretz CA. Metalloprotease domain latency protects ADAMTS13 against broad-spectrum inhibitors of metalloproteases while maintaining activity toward VWF. *Journal of Thrombosis and Haemostasis*.
143. Del Amo-Maestro L, Sagar A, Pompach P, Goulas T, Scavenius C, Ferrero DS, et al. An Integrative Structural Biology Analysis of Von Willebrand Factor Binding and Processing by ADAMTS-13 in Solution. *J Mol Biol*. 2021;433(13):166954.
144. Martin FJ, Amode MR, Aneja A, Austine-Orimoloye O, Azov AG, Barnes I, et al. Ensembl 2023. *Nucleic Acids Res*. 2023;51(D1):D933-D41.
145. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res*. 2022;50(D1):D439-D44.
146. Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods*. 2022;19(6):679-82.
147. Lyskov S, Chou FC, Conchuir SO, Der BS, Drew K, Kuroda D, et al. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). *PLoS One*. 2013;8(5):e63906.
148. The PyMOL Molecular Graphics System Online: Schrödinger, LLC.; [Version 2.0:[Available from: <http://www.pymol.org/pymol>].
149. Laskowski RA, Jablonska J, Pravda L, Varekova RS, Thornton JM. PDBsum: Structural summaries of PDB entries. *Protein Sci*. 2018;27(1):129-34.
150. Halawa M. Novel fusion tags for protein production, purification and research applications. Nottingham: University of Nottingham; 2020.
151. Crawley JT, Lam JK, Rance JB, Mollica LR, O'Donnell JS, Lane DA. Proteolytic inactivation of ADAMTS13 by thrombin and plasmin. *Blood*. 2005;105(3):1085-93.
152. UniProt C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res*. 2023;51(D1):D523-D31.
153. Lau YTK, Baytshtok V, Howard TA, Fiala BM, Johnson JLM, Carter LP, et al. Discovery and engineering of enhanced sumo protease enzymes. *Journal of Biological Chemistry*. 2018;293(34):13224-33.
154. Bunkoczi G, Read RJ. Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr D Biol Crystallogr*. 2011;67(Pt 4):303-12.
155. Tickle IJ, Flensburg C, Keller P, W. P, Sharff A, Vornrhein C, et al. STARANISO: Global Phasing Ltd, Cambridge, UK; 2018 [Available from: <https://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>].
156. Vornrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, et al. Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallographica Section D Biological Crystallography*. 2011;67(4):293-302.
157. Tischer A, Campbell JC, Machha VR, Moon-Tasson L, Benson LM, Sankaran B, et al. Mutational Constraints on Local Unfolding Inhibit the Rheological Adaptation of von Willebrand Factor. *J Biol Chem*. 2016;291(8):3848-59.

158. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 2004;60(Pt 12 Pt 1):2126-32.
159. Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. 2019;75(Pt 10):861-77.
160. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*. 2011;67(Pt 4):355-67.
161. Long F, Nicholls RA, Emsley P, Graeulis S, Merkys A, Vaitkus A, et al. AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol*. 2017;73(Pt 2):112-22.
162. Kowiel M, Brzezinski D, Porebski PJ, Shabalin IG, Jaskolski M, Minor W. Automatic recognition of ligands in electron density by machine learning. *Bioinformatics*. 2019;35(3):452-61.
163. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-12.
164. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007;372(3):774-97.
165. Scully M, Rayment R, Clark A, Westwood JP, Cranfield T, Gooding R, et al. A British Society for Haematology Guideline: Diagnosis and management of thrombotic thrombocytopenic purpura and thrombotic microangiopathies. *Br J Haematol*. 2023;203(4):546-63.
166. Joly BS, Coppo P, Veyradier A. Pediatric thrombotic thrombocytopenic purpura. *European Journal of Haematology*: Blackwell Publishing Ltd; 2018. p. 425-34.
167. Hospitals NUCL. The United Kingdom Thrombotic Thrombocytopenic Purpura (TTP) Registry [Internet] [Available from: <https://www.uclh.nhs.uk/our-services/find-service/cancer-services/blood-diseases-clinical-haematology/blood-diseases-types-and-services/red-cell-diseases/TTP/united-kingdom-thrombotic-thrombocytopenic-purpura-ttp-registry>].
168. Page EE, Kremer Hovinga JA, Terrell DR, Vesely SK, George JN. Thrombotic thrombocytopenic purpura: diagnostic criteria, clinical features, and long-term outcomes from 1995 through 2015. *Blood Adv*. 2017;1(10):590-600.
169. De Cock E, Hermans C, De Raeymaecker J, De Ceunynck K, De Maeyer B, Vandeputte N, et al. The novel ADAMTS13-p.D187H mutation impairs ADAMTS13 activity and secretion and contributes to thrombotic thrombocytopenic purpura in mice. *Journal of Thrombosis and Haemostasis*. 2015;13(2):283-92.
170. Rurali E, Banterla F, Donadelli R, Bresin E, Galbusera M, Gastoldi S, et al. ADAMTS13 secretion and residual activity among patients with congenital thrombotic thrombocytopenic purpura with and without renal impairment. *Clinical Journal of the American Society of Nephrology*. 2015;10(11):2002-12.
171. Meyer SC, Jin S, Cao W, Zheng XL, Lammler B, Hovinga JAK. Characterization of Five Homozygous ADAMTS13 Mutations in Hereditary Thrombotic Thrombocytopenic Purpura – Towards a Phenotype-Genotype Correlation? *Blood*. 2008;112(11):274-.

172. Schelpe A-S, Orlando C, Ercig B, Geeroms C, Pareyn I, Vandeputte N, et al. Child-onset thrombotic thrombocytopenic purpura caused by p.R498C and p.G259PfsX133 mutations in ADAMTS13. *European Journal of Haematology*. 2018;101(2):191-9.
173. Underwood M, Peyvandi F, Garagiola I, Machin S, Mackie I. Degradation of two novel congenital TTP ADAMTS13 mutants by the cell proteasome prevents ADAMTS13 secretion. 2016.
174. Markham-Lee Z, Morgan NV, Emsley J. Inherited ADAMTS13 mutations associated with Thrombotic Thrombocytopenic Purpura: a short review and update. *Platelets*. 2023;34(1):2138306.
175. Plaimauer B, Fuhrmann J, Mohr G, Wernhart W, Bruno K, Ferrari S, et al. Modulation of ADAMTS13 secretion and specific activity by a combination of common amino acid polymorphisms and a missense mutation. *Blood*. 2006;107(1):118-25.
176. Matsumoto M, Kokame K, Soejima K, Miura M, Hayashi S, Fujii Y, et al. Molecular characterization of ADAMTS13 gene mutations in Japanese patients with Upshaw-Schulman syndrome. *Blood*. 2004;103(4):1305-10.
177. Hommais A, Rayes J, Houllier A, Obert B, Legendre P, Veyradier A, et al. Molecular characterization of four ADAMTS13 mutations responsible for congenital thrombotic thrombocytopenic purpura (Upshaw-Schulman syndrome). *Thrombosis and Haemostasis*. 2007;98(3):593-9.
178. Zhou Z, Yeh HC, Jing H, Wang C, Tao Z, Choi H, et al. Cysteine residues in CUB-1 domain are critical for ADAMTS13 secretion and stability. *Thromb Haemost*. 2011;105(1):21-30.
179. Zhou Z, Jing H, Tao Z, Choi H, Aboulfatova K, Moake J, et al. Effects of naturally occurring mutations in CUB-1 domain on synthesis, stability, and activity of ADAMTS-13. *Thromb Res*. 2009;124(3):323-7.
180. Scully M, Gattens M, Khair K, Liesner R. The use of intermediate purity factor VIII concentrate BPL 8Y as prophylaxis and treatment in congenital thrombotic thrombocytopenic purpura. *British Journal of Haematology*. 2006;135(1):101-4.
181. Lester WA, Williams MD, Allford SL, Enayat MS, Machin SJ. Successful treatment of congenital thrombotic thrombocytopenic purpura using the intermediate purity factor VIII concentrate BPL 8Y. *Br J Haematol*. 2002;119(1):176-9.
182. Chen H, Fu A, Wang J, Wu T, Li Z, Tang J, et al. Rituximab as first-line treatment for acquired thrombotic thrombocytopenic purpura. *J Int Med Res*. 2017;45(3):1253-60.
183. Westwood JP, Thomas M, Alwan F, McDonald V, Benjamin S, Lester WA, et al. Rituximab prophylaxis to prevent thrombotic thrombocytopenic purpura relapse: outcome and evaluation of dosing regimens. *Blood Adv*. 2017;1(15):1159-66.
184. Dekimpe C, Roose E, Sakai K, Tersteeg C, De Meyer SF, Vanhoorelbeke K. Toward gene therapy for congenital thrombotic thrombocytopenic purpura. *J Thromb Haemost*. 2023;21(5):1090-9.
185. Adrienne Kitts M, Lon Phan, PhD, Minghong Ward, MS, and John Bradley Holmes, PhD. The Database of Short Genetic Variation (dbSNP). The NCBI Handbook [Internet] 2nd edition 2013.

186. Cooper DN, Ball EV, Krawczak M. The human gene mutation database. *Nucleic Acids Res.* 1998;26(1):285-7.
187. Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet.* 2017;136(6):665-77.
188. Hing ZA, Schiller T, Wu A, Hamasaki-Katagiri N, Struble EB, Russek-Cohen E, et al. Multiple in silico tools predict phenotypic manifestations in congenital thrombotic thrombocytopenic purpura. *Br J Haematol.* 2013;160(6):825-37.
189. Zhao T, Fan S, Sun L. The global carrier frequency and genetic prevalence of Upshaw-Schulman syndrome. *BMC Genom Data.* 2021;22(1):50.
190. Rim JH, Choi YJ, Gee HY. Genomic Landscape and Mutational Spectrum of ADAMTS Family Genes in Mendelian Disorders Based on Gene Evidence Review for Variant Interpretation. *Biomolecules.* 2020;10(3).
191. Rizzo C, Rizzo S, Scirè E, Di Bona D, Ingrassia C, Franco G, et al. Thrombotic thrombocytopenic purpura: a review of the literature in the light of our experience with plasma exchange. *Blood Transfus.* 2012;10(4):521-32.
192. Van Dorland HA, Taleghani MM, Sakai K, Friedman KD, George JN, Hrachovinova I, et al. The International Hereditary Thrombotic Thrombocytopenic Purpura Registry: Key findings at enrollment until 2017. *Haematologica.* 2019;104(10):2107-15.
193. Van Dorland AA, Mansouri Taleghani M, Friedman KD, George J, Hrachovinova I, Knöbl P, et al. Genotype-Phenotype Correlation in Congenital TTP: New Insights from a Multicentre Study with 121 Patients. *Blood.* 2018;132(Supplement 1):376-.
194. Wang J, Luttrell J, Zhang N, Khan S, Shi N, Wang MX, et al. Exploring Human Diseases and Biological Mechanisms by Protein Structure Prediction and Modeling. *Adv Exp Med Biol.* 2016;939:39-61.
195. Helliwell JR. New developments in crystallography: exploring its technology, methods and scope in the molecular biosciences. *Biosci Rep.* 2017;37(4).
196. Shi Y. A glimpse of structural biology through X-ray crystallography. *Cell.* 2014;159(5):995-1014.
197. Agirre J, Atanasova M, Bagdonas H, Ballard CB, Basle A, Beilsten-Edmands J, et al. The CCP4 suite: integrative software for macromolecular crystallography. *Acta Crystallogr D Struct Biol.* 2023;79(Pt 6):449-61.
198. Petri A. The allosteric activation of ADAMTS13: Imperial College London; October 2018.
199. Kobe B, Ve T, Williams SJ. Fusion-protein-assisted protein crystallization. *Acta Crystallogr F Struct Biol Commun.* 2015;71(Pt 7):861-9.
200. Ki MR, Pack SP. Fusion tags to enhance heterologous protein expression. *Appl Microbiol Biotechnol.* 2020;104(6):2411-25.
201. Garcia-Alvarez B, de Pereda JM, Calderwood DA, Ulmer TS, Critchley D, Campbell ID, et al. Structural determinants of integrin recognition by talin. *Mol Cell.* 2003;11(1):49-58.
202. Van Wart HE, Birkedal-Hansen H. The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire

- matrix metalloproteinase gene family. *Proc Natl Acad Sci U S A*. 1990;87(14):5578-82.
203. Longpre JM, Leduc R. Identification of prodomain determinants involved in ADAMTS-1 biosynthesis. *J Biol Chem*. 2004;279(32):33237-45.
204. Loechel F, Overgaard MT, Oxvig C, Albrechtsen R, Wewer UM. Regulation of human ADAM 12 protease by the prodomain. Evidence for a functional cysteine switch. *J Biol Chem*. 1999;274(19):13427-33.
205. B E, T A, J H, K W, Cpm R, K V, et al. Conformational plasticity of ADAMTS13 in hemostasis and autoimmunity. *The Journal of biological chemistry*. 2021;297(4).
206. Aponte-Santamaría C, Lippok S, Mittag JJ, Obser T, Schneppenheim R, Baldauf C, et al. Mutation G1629E Increases von Willebrand Factor Cleavage via a Cooperative Destabilization Mechanism. *Biophysical Journal*. 2017;112(1):57-65.
207. Lippok S, Radtke M, Obser T, Kleemeier L, Schneppenheim R, Budde U, et al. Shear-Induced Unfolding and Enzymatic Cleavage of Full-Length VWF Multimers. *Biophys J*. 2016;110(3):545-54.
208. Maurer S. Developing novel fusion tags to facilitate protein production and structure determination. PhD thesis, University of Nottingham: University of Nottingham; 2021.
209. Pos W, Luken BM, Sorvillo N, Kremer Hovinga JA, Voorberg J. Humoral immune response to ADAMTS13 in acquired thrombotic thrombocytopenic purpura. *J Thromb Haemost*. 2011;9(7):1285-91.
210. Jacobsen JA, Major Jourden JL, Miller MT, Cohen SM. To bind zinc or not to bind zinc: an examination of innovative approaches to improved metalloproteinase inhibition. *Biochim Biophys Acta*. 2010;1803(1):72-94.
211. Redler RL, Das J, Diaz JR, Dokholyan NV. Protein Destabilization as a Common Factor in Diverse Inherited Disorders. *J Mol Evol*. 2016;82(1):11-6.
212. Sen N, Anishchenko I, Bordin N, Sillitoe I, Velankar S, Baker D, et al. Characterizing and explaining the impact of disease-associated mutations in proteins without known structures or structural homologs. *Brief Bioinform*. 2022;23(4).
213. McPherson A, Cudney B. Optimization of crystallization conditions for biological macromolecules. *Acta Crystallogr F Struct Biol Commun*. 2014;70(Pt 11):1445-67.
214. Kangro K, Roose E, Dekimpe C, Vandenbulcke A, Graca NAG, Voorberg J, et al. Improvement of recombinant ADAMTS13 production through a more optimal signal peptide or an N-terminal fusion protein. *J Thromb Haemost*. 2022;20(10):2379-85.
215. Jankowska KI, Katneni U, Lin BC, Amarasinghe R, Phue J-N, Wu WW, et al. An Optimized Purification Design for Extracting Active ADAMTS13 from Conditioned Media. *Processes*. 2022;10(2):322.
216. Akiyama M, Nakayama D, Takeda S, Kokame K, Takagi J, Miyata T. Crystal structure and enzymatic activity of an ADAMTS-13 mutant with the East Asian-specific P475S polymorphism. *J Thromb Haemost*. 2013;11(7):1399-406.
217. Nowak AA, O'Brien HER, Henne P, Doerr A, Vanhoorelbeke K, Laffan MA, et al. ADAMTS-13 glycans and conformation-dependent activity. *J Thromb Haemost*. 2017;15(6):1155-66.

218. Kwong PD, Wyatt R, Desjardins E, Robinson J, Culp JS, Hellmig BD, et al. Probability analysis of variational crystallization and its application to gp120, the exterior envelope glycoprotein of type 1 human immunodeficiency virus (HIV-1). *J Biol Chem.* 1999;274(7):4115-23.
219. Colige AC. Challenges and solutions for purification of ADAMTS proteases: an overview. *ADAMTS Proteases: Methods and Protocols.* 2020:45-53.
220. Schelpe AS, Petri A, Roose E, Pareyn I, Deckmyn H, De Meyer SF, et al. Antibodies that conformationally activate ADAMTS13 allosterically enhance metalloprotease domain function. *Blood Advances.* 2020;4(6):1072-80.
221. DeYoung V, Singh K, Kretz CA. Mechanisms of ADAMTS13 regulation. *J Thromb Haemost.* 2022.
222. Chao LP, Einstein ER. Estimation of the molecular weight of flexible disordered proteins by exclusion chromatography. *J Chromatogr.* 1969;42(4):485-92.
223. Kokame K, Nobe Y, Kokubo Y, Okayama A, Miyata T. FRET-VWF73, a first fluorogenic substrate for ADAMTS13 assay. *Br J Haematol.* 2005;129(1):93-100.
224. Holcomb J, Spellmon N, Zhang Y, Doughan M, Li C, Yang Z. Protein crystallization: Eluding the bottleneck of X-ray crystallography. *AIMS Biophys.* 2017;4(4):557-75.
225. Schetz JA, Shankar EPN. Protein Expression in the Drosophila Schneider 2 Cell System. *Current Protocols in Neuroscience.* 2004;27(1):4.16.1-4.5.
226. Bokhove M, Sadat Al Hosseini H, Saito T, Dioguardi E, Gegenschatz-Schmid K, Nishimura K, et al. Easy mammalian expression and crystallography of maltose-binding protein-fused human proteins. *J Struct Biol.* 2016;194(1):1-7.
227. Chang VT, Crispin M, Aricescu AR, Harvey DJ, Nettleship JE, Fennelly JA, et al. Glycoprotein structural genomics: solving the glycosylation problem. *Structure.* 2007;15(3):267-73.
228. Shen X, Dojcinovic D, Baldi L, Hacker DL, Luescher IF, Wurm FM. Improved process conditions for increasing expression of MHC class II protein from a stable Drosophila S2 cell line. *Biotechnol Lett.* 2018;40(1):85-92.
229. Xie Q, Michel PO, Baldi L, Hacker DL, Zhang X, Wurm FM. TubeSpin bioreactor 50 for the high-density cultivation of Sf-9 insect cells in suspension. *Biotechnol Lett.* 2011;33(5):897-902.
230. Duggan S. Correction to: Caplacizumab: First Global Approval. *Drugs.* 2018;78(18):1955.
231. Deng W, Wang Y, Druzak SA, Healey JF, Syed AK, Lollar P, et al. A discontinuous autoinhibitory module masks the A1 domain of von Willebrand factor. *J Thromb Haemost.* 2017;15(9):1867-77.
232. Deng W, Voos KM, Colucci JK, Legan ER, Ortlund EA, Lollar P, et al. Delimiting the autoinhibitory module of von Willebrand factor. *J Thromb Haemost.* 2018;16(10):2097-105.
233. Peyvandi F, Scully M, Kremer Hovinga JA, Cataland S, Knobl P, Wu H, et al. Caplacizumab for Acquired Thrombotic Thrombocytopenic Purpura. *N Engl J Med.* 2016;374(6):511-22.
234. Arce N LR. Inhibition of von Willebrand factor through stabilization of the ristocetin-binding site. *ISTH 2022 Congress; London2022.*

235. Ward S, O'Sullivan JM, O'Donnell JS. von Willebrand factor sialylation-A critical regulator of biological function. *J Thromb Haemost.* 2019;17(7):1018-29.
236. Canis K, McKinnon TA, Nowak A, Panico M, Morris HR, Laffan M, et al. The plasma von Willebrand factor O-glycome comprises a surprising variety of structures including ABH antigens and disialosyl motifs. *J Thromb Haemost.* 2010;8(1):137-45.
237. Emsley J, Cruz M, Handin R, Liddington R. Crystal structure of the von Willebrand Factor A1 domain and implications for the binding of platelet glycoprotein Ib. *J Biol Chem.* 1998;273(17):10396-401.
238. Voos KM, Cao W, Arce NA, Legan ER, Wang Y, Shajahan A, et al. Desialylation of O-glycans activates von Willebrand factor by destabilizing its autoinhibitory module. *J Thromb Haemost.* 2022;20(1):196-207.
239. De Luca M, Facey DA, Favaloro EJ, Hertzberg MS, Whisstock JC, McNally T, et al. Structure and function of the von Willebrand factor A1 domain: analysis with monoclonal antibodies reveals distinct binding sites involved in recognition of the platelet membrane glycoprotein Ib-IX-V complex and ristocetin-dependent activation. *Blood.* 2000;95(1):164-72.
240. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *The FASEB Journal.* 2000;14(2):231-41.
241. George RA, Heringa J. An analysis of protein domain linkers: their classification and role in protein folding. *Protein Engineering, Design and Selection.* 2002;15(11):871-9.
242. O'Sullivan B, Smith B, Baramidze G. Recent advances for seeding a crystallization process. *Mettler Toledo Auto-Chem, Columbia.* 2012.
243. He Y, Gao Z, Zhang T, Sun J, Ma Y, Tian N, et al. Seeding Techniques and Optimization of Solution Crystallization Processes. *Organic Process Research & Development.* 2020;24(10):1839-49.
244. Obmolova G, Malia TJ, Teplyakov A, Sweet RW, Gilliland GL. Protein crystallization with microseed matrix screening: application to human germline antibody Fabs. *Acta Crystallogr F Struct Biol Commun.* 2014;70(Pt 8):1107-15.
245. Uchida T, Wada H, Mizutani M, Iwashita M, Ishihara H, Shibano T, et al. Identification of novel mutations in ADAMTS13 in an adult patient with congenital thrombotic thrombocytopenic purpura. *Blood.* 2004;104:2081-3.
246. Grall M, Azoulay E, Galicier L, Provot F, Wynnckel A, Poullin P, et al. Thrombotic thrombocytopenic purpura misdiagnosed as autoimmune cytopenia: Causes of diagnostic errors and consequence on outcome. Experience of the French thrombotic microangiopathies reference centre. *Am J Hematol.* 2017;92(4):381-7.
247. Niazi SK, Mariam Z, Paracha RZ. Limitations of Protein Structure Prediction Algorithms in Therapeutic Protein Development. *BioMedInformatics.* 2024;4(1):98-112.
248. Terwilliger TC, Afonine PV, Liebschner D, Croll TI, McCoy AJ, Oeffner RD, et al. Accelerating crystal structure determination with iterative AlphaFold prediction. *Acta Crystallogr D Struct Biol.* 2023;79(Pt 3):234-44.
249. Pace NJ, Weerapana E. Zinc-binding cysteines: diverse functions and structural motifs. *Biomolecules.* 2014;4(2):419-34.

250. Goshua G, Sinha P, Hendrickson JE, Tormey C, Bendapudi PK, Lee AI. Cost effectiveness of caplacizumab in acquired thrombotic thrombocytopenic purpura. *Blood*. 2021;137(7):969-76.
251. Ulrichs H, Silence K, Schoolmeester A, de Jaegere P, Rossenu S, Roodt J, et al. Antithrombotic drug candidate ALX-0081 shows superior preclinical efficacy and safety compared with currently marketed antiplatelet drugs. *Blood*. 2011;118(3):757-65.
252. Prestegard JH. A perspective on the PDB's impact on the field of glycobiology. *J Biol Chem*. 2021;296:100556.
253. Scully M, Knöbl P, Kentouche K, Rice L, Windyga J, Schneppenheim R, et al. Clinical Trials and Observations: Recombinant ADAMTS-13: first-in-human pharmacokinetics and safety in congenital thrombotic thrombocytopenic purpura. *Blood*. 2017;130(19):2055-.
254. Jain N, Marquez C, Martell L. Recombinant ADAMTS13 for Patients with Severe Congenital Thrombotic Thrombocytopenic Purpura: Design of a Phase 3b Open-Label Continuation Study of Prophylactic and on-Demand Treatment. *Blood*. 2021;138(Supplement 1):4252-.