UNIVERSITÄT BONN

Structure and functional characterization of the B-domain and the full-length coagulation factor VIII

Samhitha Urs Ramaraje Urs

Structure and functional characterization of the B-domain and the full-length coagulation factor VIII

……

DISSERTATION

zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Samhitha Urs Ramaraje Urs

aus

Mysuru, Indien

Bonn, 2025

Aus dem Institut für experimentelle Hämatologie und Transfusionsmedizin der Rheinischen Friedrich-Wilhelms-Universität Bonn

………………………………………………………………………………………………

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer:	PD. Dr. Arijit Biswas
Gutachterin:	Prof. Dr. Diana Imhof

Tag der mündlichen Prüfung: 11.08.2025

Erscheinungsjahr: 2025

विद्या ददाति विनयं, विनयाद्याति पात्रताम्।
पात्रत्वाद्धनमाप्नोति धनाद्धर्मं ततः सुखम्॥

“Knowledge brings humility, humility brings worthiness, worthiness leads to wealth/enrichment, enrichment leads to righteousness, and righteousness brings happiness.”

- **‘Hitopadesha’**- Pandita Narayana.

Owing everything I do, achieve, and aspire to be to the very essence of my soul and existence - my beloved Grandparents, my Partner, my Parents, and my Brother.
Forever humbled and Grateful!

LIST OF PUBLICATIONS:

1. **Samhitha Urs Ramaraje Urs**[1], Jean-Luc Pellequer[2], Jean-Marie Teulon[2], Boxue Ma[3], Deniz Ugular[3], Sneha Singh[1], Jens Müller[1], Simone Gasper[1], Anna Pepanian[3], Diana Imhoff[3] , Johannes Oldenburg[1], Arijit Biswas[1] (2025). Decoding the B-Domain: Exploring Structural and Functional Dynamics in Full-Length Factor VIII - Manuscript under preparation.

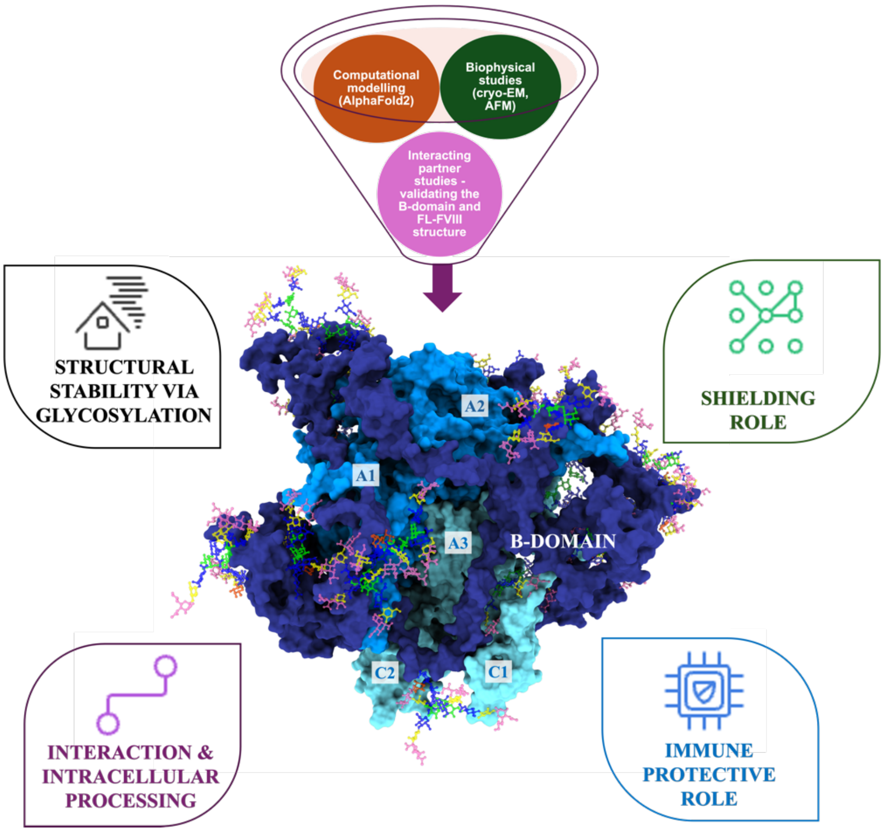
2. Alua Kuanyshbek[1]*, Hamideh Yadegari[1]*, Jens Müller[1], Nasim Shahidi Hamedani[1], **Samhitha Urs Ramaraje Urs**[1], Johannes Oldenburg[1] (2025). Interplay between Circulating von Willebrand Factor and Neutrophils: Implications on Inflammation, Neutrophil Function, and von Willebrand Factor clearance - *Hematologica*
   DOI: https://doi.org/10.3324/haematol.2023.284919

3. Sneha Singh[1]*, Gregor Hagelueken[2]*, Deniz Ugurlar[3], **Samhitha Urs Ramaraje Urs[1]**, Amit Sharma[4], Manoranjan Mahapatra[4], Friedel Drepper[5], Diana Imhof[6], Pitter F. Huesgen[5], Johannes Oldenburg[1], Matthias Geyer[2], Arijit Biswas[1] (2024). Cryo-EM structure of the human native plasma coagulation Factor XIII complex - *Blood*
   DOI: https://doi.org/10.1182/blood.2024025369

4. Behnaz Pezeshkpoor PhD[1], Nadja Sereda M.Sc[1], Janine Becker-Gotot PhD[2], Ann-Cristin Berkemeier[1], Isabell Matuschek, M.Sc[1,] Jens Müller PhD[1], **Samhitha Urs Ramaraje Urs M.Sc.[1]**, Sneha Singh PhD[1], Natascha, Marquardt MD[1,3], J. Oldenburg MD[1,3] (2024). Comprehensive analysis of Neutralizing Anti-Emicizumab Antibodies on Drug Efficacy in Acquired Hemophilia A - *Journal of Thrombosis and Haemostasis (JTH)*
   DOI: https://doi.org/10.1016/j.jtha.2024.10.003

5. Haroon Javed[1], Sneha Singh[1], **Samhitha Urs Ramaraje Urs[1]**, Johannes Oldenburg[1], Arijit Biswas[1] **(2022).** Genetic landscape in coagulation factor XIII associated defects - Advances in coagulation and beyond - *Blood Reviews*
   DOI: https://doi.org/10.1016/j.blre.2022.101032

GRAPHICAL ABSTRACT



ABSTRACT

Factor VIII (FVIII) is a large glycoprotein that serves as a non-enzymatic cofactor for activated Factor IX, facilitating the proteolytic activation of Factor X in the coagulation cascade. Deficiency of FVIII leads to Hemophilia-A (HA), a bleeding disorder. Synthesized as a single-chain protein (~300 kDa) by liver sinusoidal endothelial cells, FVIII undergoes processing to yield a mature protein with the domain organization A1-a1-A2-a2-B-a3-C1-C2. In circulation, FVIII exists as a heterodimer comprising a heavy chain (A1-A2-B) and a light chain (A3-C1-C2) and stabilized by divalent metal ions such as copper (Cu^{2+}) and calcium (Ca^{2+}). The B-domain, of approximately 1,000 amino acids, is the largest and most heavily glycosylated region of FVIII. Since the B-domain is cleaved off during the activation of FVIII, it is dispensable for cofactor activity as evidenced in recombinant FVIII products lacking B domain. Although not critical to FVIII activity, the B-domain plays a key but not fully understood role in FVIII synthesis, secretion, and intracellular stability. This B-domain deleted FVIII (BDD-FVIII), forms the basis for many recombinant FVIII products used in HA replacement therapies. The B-domain remains structurally uncharacterized, which poses a significant challenge to the complete understanding of FVIII biology. To bridge this gap, we employed an integrative hybrid approach to develop detailed structural models of glycosylated and furin-cleaved full-length FVIII (*Gly/FC-FL-FVIII*), including the B-domain. Leveraging the Nobel-recognized AlphaFold2 (AF2) predictions, we refined these models and validated them with biophysical data from atomic force microscopy (AFM) and cryo-electron microscopy (cryo-EM) analyses of recombinant and plasma-derived full-length FVIII proteins. We further extended our studies by modeling a full-length FVIII-von Willebrand Factor (vWF) complex, incorporating both the B-domain and D′-D3 domains (*FL-FVIII-vWF-D'-D3*), to explore the B-domain's interactions during intracellular processing. Our findings reveal that the B-domain functions predominantly as an independent domain that encircles the FVIII core. Structural models indicate that glycosylations contribute to protein stability and proper folding. Cryo-EM and AFM studies also revealed a globular structure of FL-FVIII, with vWF in a tail-like configuration. Structural models of the *FL-FVIII-vWF-D'-D3* complex suggest that furin-mediated cleavages allow the B-domain to unfold and facilitate critical interactions with vWF and other partners. Additionally, the B-domain appears to shield key FVIII epitopes, reducing immune recognition and mitigating premature activation. Further computational studies shed light on the B-domain's contribution to FVIII secretion, stability, and immune evasion. This study introduces a robust structural-functional model of the B-domain and FL-FVIII, offering novel insights into the B-domain's role in FVIII biology. These findings provide a valuable foundation for advancing FVIII-based therapeutic strategies, addressing challenges in HA treatment, and enhancing our understanding of FVIII biology.

ZUSAMMENFASSUNG

Faktor VIII (FVIII) ist ein großes glykoprotein, das als nicht-enzymatischer kofaktor für aktivierten Faktor IX dient und die proteolytische Aktivierung von Faktor X in der gerinnungskaskade erleichtert. Ein mangel an FVIII führt zu Hämophilie-A (HA). FVIII wird von den sinusoidalen endothelzellen der leber als einkettiges Protein (~300 kDa) synthetisiert und zu einem reifen Protein mit der Domänenorganisation A1-a1-A2-a2-B-a3-C1-C2 verarbeitet. Im Blutkreislauf liegt FVIII als Heterodimer vor, das eine schwere kette (A1-A2-B) und eine leichte kette (A3-C1-C2) umfasst und durch zweiwertige metallionen wie Kupfer (Cu^{2+}) und Calcium (Ca^{2+}) stabilisiert wird. Die B-Domäne mit etwa 1.000 aminosäuren ist die größte und am stärksten glykosylierte region von FVIII. Die B-Domäne wird zwar gespalten und ist für die kofaktor Aktivität entbehrlich, ist aber bekanntermaßen für die FVIII-Synthese, die Sekretion und die intrazelluläre Stabilität entscheidend. Diese aktivierte Form von FVIII, auch B-domain deleted FVIII (BDD-FVIII) genannt, bildet die grundlage für viele rekombinante FVIII-produkte, die in

HA-ersatztherapien verwendet werden. Die B-Domäne ist strukturell noch nicht charakterisiert, was eine große herausforderung für ein tieferes Verständnis der Biologie von FVIII darstellt. Daher haben wir einen integrativen hybridansatz gewählt, um diese Lücken zu schließen und detaillierte Strukturmodelle von glykosyliertem und Furin-gespaltenem FVIII in voller Länge (*Gly/FC-FL-FVIII*), einschließlich der B-Domäne, zu entwickeln. Auf der grundlage der von Nobelpreisträger AlphaFold2 (AF2) erstellten Vorhersagen haben wir diese Modelle verfeinert und mit biophysikalischen Daten aus der Rasterkraftmikroskopie (AFM) und der Kryo-elektronenmikroskopie (kryo-EM) von rekombinanten und aus Plasma gewonnenen FVIII-Proteinen validiert. Wir haben unsere Studien erweitert, indem wir einen FVIII-von Willebrand-Faktor (vWF)-komplex in voller Länge modelliert haben, der sowohl die B-Domäne als auch die D′-D3-Domäne (*FL-FVIII-vWF-D'-D3*) enthält, um die Interaktionen der B-Domäne während der intrazellulären Verarbeitung zu untersuchen. Unsere Ergebnisse zeigen, dass die B-Domäne überwiegend als unabhängige domäne fungiert, die den FVIII-Kern umschließt. Strukturelle Modelle deuten darauf hin, dass glykosylierungen zur Proteinstabilität und zur korrekten Faltung beitragen. Kryo-EM- und AFM-studien ergaben außerdem eine kugelförmige Struktur von FL-FVIII, mit vWF in einer schwanzartigen konfiguration. Strukturmodelle des *FL-FVIII-vWF-D'-D3*-Komplexes legen nahe, dass Furin-vermittelte Spaltungen die Entfaltung der B-Domäne ermöglichen und kritische Interaktionen mit vWF und anderen Partnern erleichtern. Darüber hinaus scheint die B-Domäne wichtige FVIII-Epitope abzuschirmen, wodurch die Erkennung durch das Immunsystem verringert und die vorzeitige Aktivierung abgeschwächt wird. Weitere computergestützte Studien geben Aufschluss über den beitrag der B-Domäne zur FVIII-Sekretion, Stabilität und Immunabwehr. Diese Studie stellt ein robustes strukturell-funktionelles Modell der B-Domäne und FL-FVIII vor und bietet neue Einblicke in die Rolle der B-Domäne in der FVIII-biologie. Diese Erkenntnisse bilden eine wertvolle Grundlage für die Weiterentwicklung von FVIII-basierten therapeutischen Strategien, die Bewältigung von Herausforderungen bei der HA-Behandlung und die Verbesserung unseres Verständnisses der FVIII-Biologie.

TABLE OF CONTENTS:

LIST OF ABBREVIATIONS:

° C	Grad Celsius
µl	Microliter
µg	Microgram
µg/µl	Microgram/Microliter
aa	Amino acids
A/Ala	Alanine
ACN	Acetonitrile
(AcUr)$UO_2(CH_3COO)_2.2H_2O$	Uranyl Acetate
Ag	Antigen
AFM	Atomic Force Microscopy
APC	Activated Protein C
aPTT	Activated partial thromboplastin time
approx.	Approximately
R/Arg	Arginine
D/Asp	Aspartic Acid
N/Asn	Asparagine
ASPGR	Asialoglycoprotein receptor
Ab	Antibody
APBS	Adaptive Poisson-Boltzmann Solver
BHK	Baby Hamster Kidney cells
BDD-FVIII	B-domain deleted Factor VIII
bp	Base Pairs
BSA	Bovine Serum Albumin
Ca	Calcium
$CaCl_2$	Calcium chloride
Cß	C-beta
CCD	Charge-Coupled Diode
CCS	Cross-Correlation Score
CD	Circular Dichroism
Co-IP	Co-Immunoprecipitation
COOT	Crystallographic Object-Oriented Toolkit
CTF	Contrast Transfer Function
CV	Column Volume
CHAMP	CDC Hemophilia A Mutation Projects
CHO	Chinese Hamster Ovary Cels
CO_2	Carbon dioxide
C/Cys	Cysteine
Cryo-EM	cryogenic-electron microscopy
Da	Dalton

DED	Direct electron detectors
DNA	Deoxyribonucleic acid
DPBS	Diphosphate buffered saline
DSSP	Define Secondary Structure of Proteins
DTT	Dithiothreitol
EAHAD	European Association for Haemophilia and Allied Disorders
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic Reticulum
E.coli	Escherichia coli
EGF	Epidermal growth factor-like domain
ERGIC	Endoplasmic Reticulum Golgi intermediate compartment
FIIa	Thrombin
FIXa	Activated Factor IX
FVa	Activated Factor V
FVIIIa	Activated Factor VIII
FVIII:Ag	Factor VIII-Antigen
FF	Force Field
FIX	Factor IX
FV	Factor V
FVIII	Factor VIII
FL-FVIII	Full-length Factor VIII
FSC	Fourier Shell Correlation
FVIII: Ac	Factor VIII-Activity
FX	Factor X
g	Gravity of earth [meters per square second]
E/Glu	Glutamic acid
Q/Gln	Glutamine
G/Gly	Glycine
H/His	Histidine
h	Hour
HCl	Hydrogen chloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High-performance liquid chromatography
HA	Haemophilia A
HEK	Human Embryonic Kidney Cels
H/His	Histidine
H_2SO_4	Sulfuric acid
I/Ile	Isoleucine
IAC	Immuno Affinity Chromatography

IEX	Ion Exchange Chromatography
IU	International Units
keV	kiloelectronvolt
Kb	Kilobases
L/Leu	Leucine
LRP	Low-density lipoprotein-related receptor protein
LDL	Low-density Lipoprotein
LMAN1	Mannose-binding lectin 1
K/Lys	Lysine
MD	Molecular Dynamics
min	Minutes
mg/ml	milligram/millileter
mm	Millimeter
MCFD2	Multiple coagulation factor deficiency protein
ml	Milliliter
MW	Molecular weight
M/Met	Methionine
MM	Multimer
MWCO	Molecular Weight Cut Off
$Na_4O_{40}SiW_{12}$	Sodium Silico Tungstate
NH4HCO3	Ammonium bicarbonate
ng	Nanogramm
nM	Nanomolar
NaCl	Sodium Chloride
NPP	Normal Pool Plasma
NMR	Nuclear Magnetic Resonance
nm	Nanometer
PACE/Furin	Paired basic Amino acid Cleaving Enzyme
pdFVIII	Plasma-derived Factor VIII protein
pg	Picogram
P/Pro	Proline
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
F/Phe	Phenylalanine
PBS	Phosphate Buffer Saline
ps	Picoseconds
P/Pro	Proline
PVDF	Polyvinylidene Fluoride
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
rFVIII	Recombinant Factor VIII protein
rpm	Revolutions per minute
RT	Room Temperature
SASA	Solvent surface area accessibility
s	Seconds
SDS	Sodium dodecyl sulfate
S/Ser	Serine
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel electrophoresis
SP	Signal Peptide
SEC	Size Exclusion Chromatography
SD	Solvent/Detergent Treatment
SSC	Secondary Structure Content
SPA	Single Particle Analysis
TAE	TRIS base, acetic acid, and EDTA.
TE	TRIS and EDTA
T/Thr	Threonine
TRIS-HCl	Tris base with hydrochloric acid
Y/Tyr	Tyrosine
TAFI	Thrombin Activated Fibrinolysis Inhibitor
TEM	Transmission Electron Microscopy
TPE	Total Potential Energy
TF	Tissue Factor
TRIS	Tris(2,3-dibromopropyl) phosphate
W/Trp	Tryptophan
UF	Ultrafiltration
UV	Ultraviolet Light
V	Volt
VLDL	very low-density Lipoprotein
V/Val	Valine
vWF	von Willebrand Factor
XL-MS	Cross-linking mass spectrometry
XRC	X-ray Crystallography
YASARA	Yet Another Scientific Reality Application

CHAPTER 1 - INTRODUCTION

1.1 The coagulation cascade

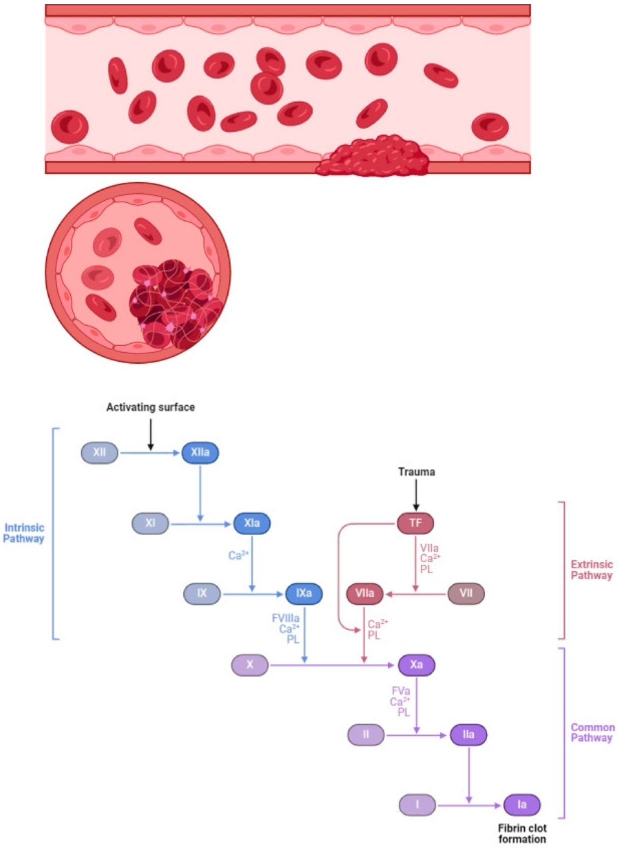


Figure 1. Overview of the coagulation cascade.

Hemostasis is a vital mammalian process that maintains blood flow and vessel integrity. Upon vascular injury, a platelet-fibrin clot forms at the site to prevent blood loss, while procoagulant substances remain localized to avoid widespread clotting. This is achieved through the organization of coagulation reactions on specific cell surfaces. [1-3] Haemostasis comprises two main components: platelet-derived and plasmatic. Injury triggers the release of tissue cells and subendothelial matrix proteins, exposing collagen that binds von Willebrand factor (vWF). Structural changes in vWF enable platelet adhesion via the glycoprotein Ib-IX complex, forming a primary platelet plug. This activation alters platelet shape, releases coagulation factors, and promotes further platelet recruitment. Plasmatic hemostasis involves a self-amplifying enzyme cascade with initiation, amplification, propagation, and termination phases. [2,4-6] Coagulation factors, designated by Roman numerals, function as vitamin K-dependent serine proteases with cofactors, assembling on cell membranes in a calcium-dependent manner to enhance efficiency. The initiation phase occurs on tissue factor (TF)-bearing cells, where TF forms a complex with activated Factor VII (FVIIa), initiating activation of Factors IX (FIX) and X (FX). FXa binds Factor Va (FVa) to form the prothrombinase complex, generating small amounts of thrombin (FIIa). In the amplification phase, thrombin activates platelets, Factor VIII (FVIII), and Factor V (FV), enhancing procoagulant activity. During propagation, activated FIX (FIXa) binds activated FVIII (FVIIIa) on activated platelets, forming the intrinsic tenase complex, which activates FX. This leads to a burst of thrombin production, further amplifying coagulation. Thrombin also converts fibrinogen to fibrin, stabilizing the clot through Factor XIII (FXIII)-mediated cross-linking (Figure 1). [3,6,7] The TF and intrinsic pathways operate on different surfaces but are essential for hemostasis. Termination mechanisms regulate thrombin production, with Tissue Factor Pathway Inhibitor (TFPI) inhibiting the TF complex, and antithrombin targeting FXa and thrombin. Thrombomodulin-bound thrombin activates protein C, which, with protein S, degrades FVa and FVIIIa to prevent excessive clotting. Finally, fibrinolysis dissolves clots via plasmin, which degrades fibrin into fragments cleared by proteases or excreted by the liver and kidneys. [4,5,8-10]

1.2 Coagulation Factor VIII

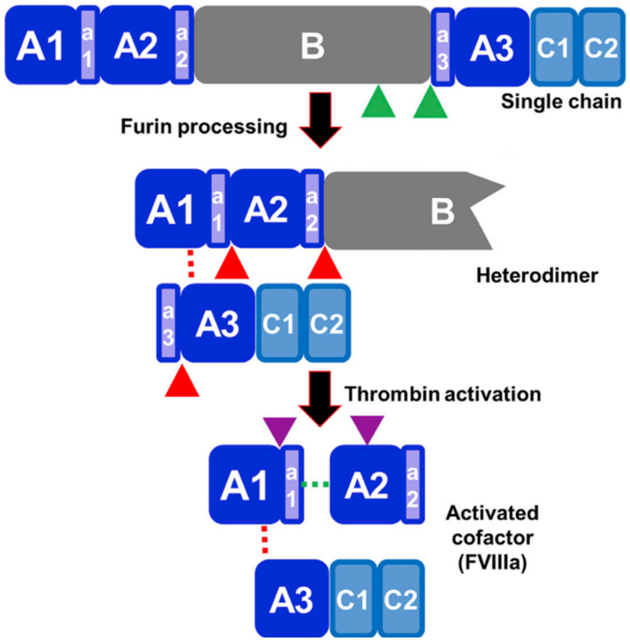


Figure 2. Domain structure and processing of FVIII protein.

The F8 gene, which encodes Factor VIII, is located on the long arm of the X chromosome at position Xq28, approximately 1,000 kb from the telomeric end. It spans a large genomic region of 186 kilobases (kb), representing around 0.1% of the total X chromosome. The gene is oriented with its 5' untranslated region (UTR) facing the telomere, and transcription proceeds toward the centromere. The F8 gene comprises 26 exons, which code for a 19-amino acid (aa) signal peptide and a mature protein consisting of 2,332 amino acids. Exon sizes range from 69 to 313 base pairs (bp), except for exon 26, which contains a 1,958 bp untranslated region, and the large exon 14, spanning 3,106 bp and accounting for about 40% of the coding sequence. The introns vary significantly in size, from 0.2 kb (intron 17) to 32 kb (intron 22). The F8 mRNA is approximately 9 kb in length, while the complementary DNA (cDNA) consists of 7,053 bp. The mRNA translates into a 2,351-aa precursor protein. Transcription initiates 170 bp upstream of the start codon (ATG, coding for methionine). A GATAAA motif, resembling a TATA box, is present 30 bp upstream of the transcription start site but has a minimal effect on promoter activity. The promoter region extends from -270 to -64 bp and contains regulatory elements specific to hepatocyte expression. Key transcription factors that bind to this region include hepatocyte nuclear factor 1 (HNF-1), D-site binding protein (DBP), and CCAAT/enhancer-binding proteins (C/EBPα and C/EBPβ). Additionally, a potential repressor element is located between -740 and -1,002 bp. The final exon, exon 26, features an extensive 1.8 kb untranslated region, contributing to post-transcriptional regulation. [11,12,13-15] FVIII is produced as a 2,332 residue glycoprotein from endothelial cells and activated platelets, functioning as an essential cofactor in the intrinsic aspect of the blood coagulation cascade. FVIII is synthesized with the domain structure A1-a1-A2-a2-B-a3-A3-C1-C2 with each A-domain containing a highly acid region (a1-, a2-, and a3-domains) heavily clustered with aspartate and glutamate residues. [16-18] The B-domain is highly glycosylated, containing most of the protein's Asn-linked carbohydrate chains. Despite its large size (~700 amino acids), the B-domain is not essential for FVIII activity. Its primary role is known to lie in intracellular processing and providing cleavage sites for secretion and thrombin-mediated activation. Initial processing of FVIII cleaves away approximately 300 residues of the B-domain, generating a heterodimer that circulates bound with its carrier protein, vWF. Heterodimeric FVIII consists of a heavy chain - A1-a1-A2-a2-B - HC $_{(A1\text{-}a1\text{-}A2\text{-}a2\text{-}B)}$ with differing lengths of the B-domain due to limited proteolysis, and a light chain - A3-a3-C1-C2 - LC $_{(a3\text{-}A3\text{-}C1\text{-}C2)}$ containing the phospholipid binding region. [19-24] vWF interacts with the FVIII light chain to protect FVIII from premature inactivation during normal hemostasis, but the exact binding interaction between FVIII and vWF is not completely understood. During vascular injury, thrombin proteolytically activates FVIII by cleaving FVIII at residues near A1-A2, A2-B, and A3-domains. Post-activation, FVIIIa is released from vWF as a heterotrimer of A1/A2/A3-C1-C2 and binds to activated platelet surfaces with FIXa to form the intrinsic Xase complex. The C-domain domains play crucial roles in its structural integrity and function in coagulation. The C1-domain enhances vWF binding when combined with C2. The C2-domain also has binding sites for vWF, thrombin, and FXa, with mutagenesis studies showing that alterations in key hydrophobic residues reduce inhibitor binding (Figure 2). [25-28,29,30]

1.3 FVIII structures and homology

Computer-aided structural analysis of protein and nucleic acid sequence databases for FV and FVIII has revealed statistical homology between these factors and human ceruloplasmin. The A-domains of FVIII exhibit homology with each other and with the copper-binding protein ceruloplasmin (specifically residues 1-77, 400-433, and 683-191), suggesting their potential role in metal ion binding. Furthermore, FVIII shares 30-40 % homology with FV, with both factors containing central B-domains rich in glycosylation sites that are released during thrombin activation. However, the B-domains of FV and FVIII show only 14 % homology with one another. Interestingly, the C-domains of FVIII demonstrate 20 % homology with phospholipid-binding proteins such as the lipid-binding lectin Discoidin-I, human and murine milk fat globule proteins, and a putative neuronal cell adhesion molecule from *Xenopus laevis*. While the functional significance of the B-domains remains unclear, they are heavily glycosylated in both FV and FVIII and are released upon cofactor activation. Despite the A- and C-domains of FV and FVIII being 40% identical, the B-domains are distinct.

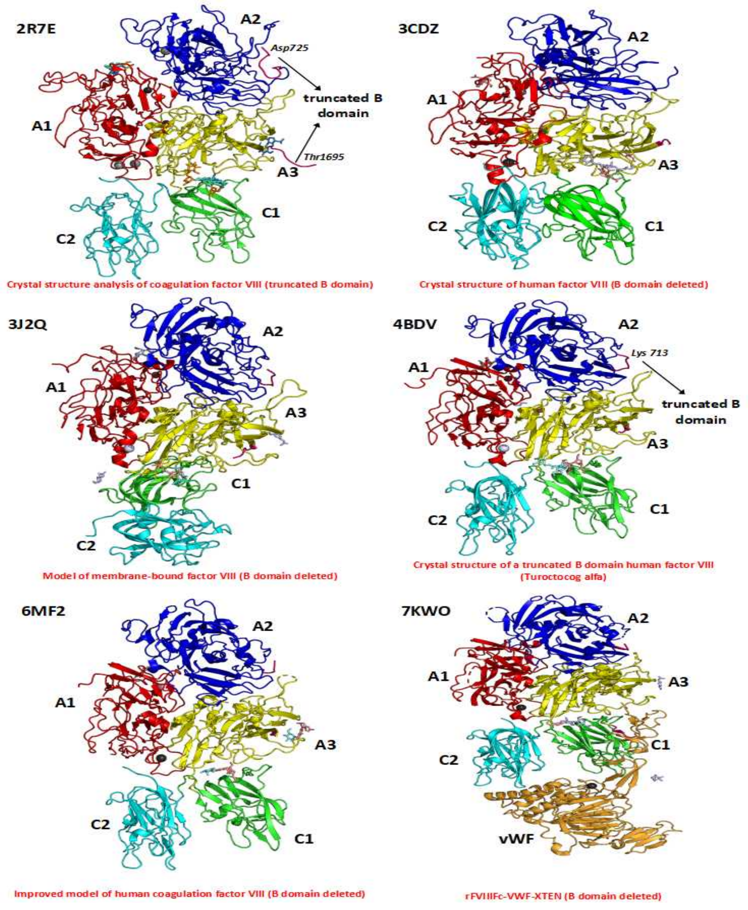


Figure 3. Summary of all the PDB deposited structures of FVIII protein showing specific domains. Red: A1, Blue: A2, Yellow: A3, Green: C1, and Cyan blue: C2, Golden yellow: VWF 2R7E- The structures are Crystal structure analysis of coagulation factor VIII (truncated B-domain), 3CDZ- Crystal structure of human factor VIII (BDD), 3J2Q- Crystal structure of human factor VIII (BDD), 4BDV- Crystal structure of a truncated B-domain human factor VIII (Turoctocog alfa), 6MF2- Improved model of human coagulation factor VIII (BDD), 7KWO- rFVIIIFc-VWF-XTEN (BDD)

However, both B-domains possess a high density of asparagine-linked oligosaccharides and are encoded on unusually large single exons, suggesting evolutionary conservation and a shared origin. Further homology studies suggest that the B-domains of FV and FVIII likely evolved through gene duplication and divergence, acquiring heavy glycosylation to maintain conserved structural features while their functional roles remain largely unexplored. This evolutionary perspective also offers valuable insights for inhibitor studies and the identification of interaction sites for ligands and macromolecules. Furthermore, the dimer structure of bovine FVa-A1/A3-C1-C2, obtained through Protein C cleavage, provides a superior model for understanding domain interactions in activated FVIIIa. Recent structural analyses of FV and FVa reveal the dynamic nature of the B-domain and its role in maintaining the assembly of heavy and light chains displayed the removal of the B-domain during FV activation increases the disorder within this assembly. By extension, a similar hypothesis can be extended to FVIII, where the B-domain likely plays a comparable regulatory role in preserving the structural integrity of the activated cofactor. These findings underscore the importance of B-domain dynamics in modulating the activity and structural stability of both FV and FVIII. High-resolution crystallographic structures of the FVIII C2-domain have elucidated hydrophobic and electrostatic interactions critical for membrane binding. A model of FVIII integrating the A- and C-domains has been developed using inter-domain orientation data from 2D crystal structures of FVIII bound to phospholipids at a resolution of 15 Å[2]. Currently, there are more than six crystal structures deposited in the RCSB PDB database (https://www.rcsb.org/), which are BDD-FVIII or contain truncated B-domain (Figure 3).[26,30-34]

1.4 FVIII biology and roles of B-domain

1.4.1 FVIII - Synthesis and secretion

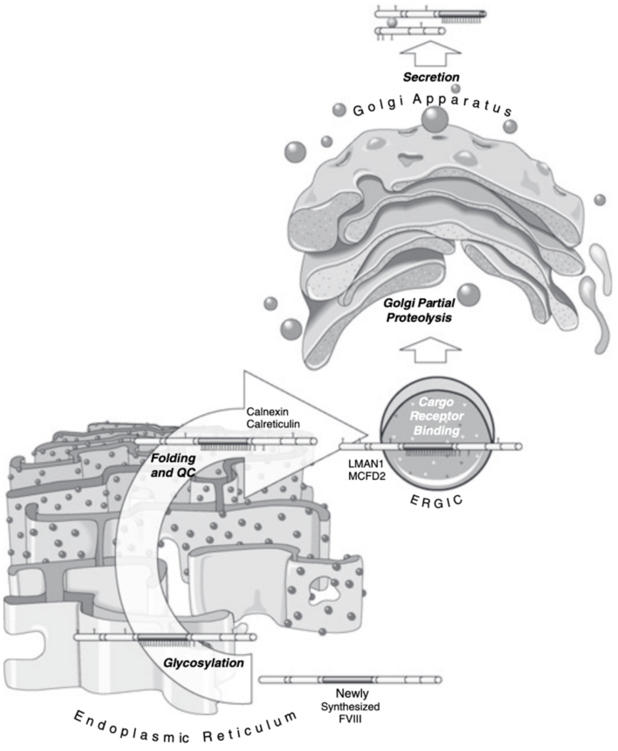


Figure 4. Overview of intracellular processing and secretion of FVIII protein.

Produced primarily in hepatocytes and LSECs, synthesis of FVIII has been demonstrated in many tissue types, including liver, spleen, kidney, and lymphatic tissue. In heterologous systems, FVIII secretion is less efficient due to impaired intracellular trafficking from the ER to the Golgi apparatus. Post-translational modifications occur co-translationally, with the B-domain serving as a key modification site. This region features the sulfation of six Y residues in acidic regions, O-linked glycosylation of S/T residues, and the addition of 14 branched high-mannose chains to N residues (N-X-S/T motifs). High-mannose glycans are critical for interactions with ER chaperones like calnexin/calreticulin (CNX/CRT) and BiP, which assist in protein folding, assembly, and quality control. The CNX/CRT cycle involves the sequential trimming of glucose residues by glucosidase I and II, facilitating proper folding. The B-domain is integral to QC and secretion, as demonstrated by the reduced efficiency following glycosylation inhibition.[35] Once correctly folded in the ER, FVIII is transported via COPII-coated vesicles to the ER-Golgi intermediate compartment (ERGIC), where mannose residues on the B-domain bind to the LMAN1/MCFD2 receptor complex. This complex, comprising ERGIC-53 (LMAN1) and MCFD2, facilitates cargo sorting in a calcium-dependent manner, with all components co-localizing in the Golgi. Proteolytic processing in the Golgi by PACE/Furin cleaves FVIII at residues 1313 and 1648 (R-X-X-R motifs), yielding a heterodimer consisting of an HC (A1-a1-A2-a2-B) and an LC (a3-A3-C1-C2). The B-domain, prone to proteolysis, forms a rod-like structure, and variability in HC length is attributed to limited C-terminal processing. In circulation, FVIII binds to the vWF, which stabilizes the complex by interacting with the LC $_{(a3\text{-}A3\text{-}C1\text{-}C2)}$ rather than the B-domain (Figure 4). Thrombin activation converts this complex into an active heterotrimer. Although FVIII can be secreted without the B-domain, efficiency is shown to decline significantly. Commercial BDD-FVIII products, like ReFacto-AF®, retain a minimal B-domain segment (S743-Q1638). Studies in COS-1 and HeLa cells show that rBDD-FVIII has reduced binding to LMAN1/MCFD2, suggesting the B-domain is crucial for this interaction. Patients with LMAN1/MCFD2 mutations exhibit combined FV and FVIII deficiencies, with protein levels at 5-30% of normal. While the B-domain is important for FVIII secretion, further research is needed to fully elucidate its role in secretion efficiency. [16,18,24,34,35]

1.4.2 FVIII - Activation.

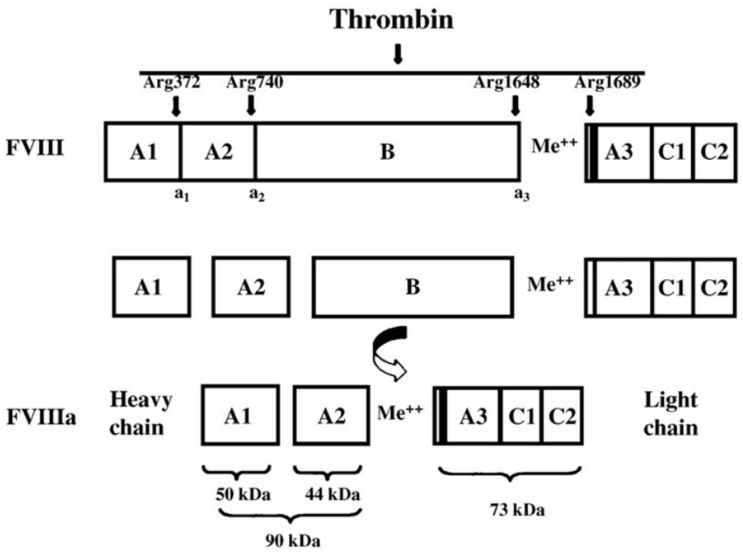


Figure 5. Overview of thrombin activation of FVIII protein.

Circulating as a pro-cofactor, FVIII undergoes proteolytic processing between B- and A3-domains by Thrombin, resulting in the formation of a heterodimer with a variable HC $_{(A1\text{-}a1\text{-}A2\text{-}a2\text{-}B)}$ (80-210 kDa) and LC $_{(a3\text{-}A3\text{-}C1\text{-}C2)}$ further processed into 73kDa, 50kDa, and 43 kDa fragments. Thrombin binds to the sulfonated acidic peptides a1, a2, and a3 and cleaves peptide bonds after these Arg residues. The proteolytic processing of the FVIII is associated with the cleavage at R740 (A2-B), in HC $_{(A1\text{-}a1\text{-}A2\text{-}a2\text{-}B)}$, produces a uniform fragment of 90kDa, further followed by cleavage at R372 (A1-A2) to produce 50 and 43kDa peptide and cleavage near N-terminus at R1689 produce 73kDa LC $_{(a3\text{-}A3\text{-}C1\text{-}C2)}$ fragments constituting the FVIIIa (Figure 5). In vitro mutagenesis studies of these R residues show that A2-B cleavage is not essential for activity, whereas those between A1- and A2- (in a1) and a3- and A3-domains are necessary for detectable clotting activity. Hemophilic mutations of these R residues have also been shown to immensely reduce the FVIII coagulant activity. Both pdFL-FVIII and rFL-FVIII are reported with similar thrombin activation, though rBDD-FVIII exhibited altered cleavage patterns with no influence by the glycosylations. Thrombin is shown to bind to the A1-domain with high affinity and the A2-domain with lower affinity. Another potential procoagulant activator of FVIII is FX which cleaves in the presence of Ca^{2+} and phospholipids. FXa-activated

FVIII shows variable activity compared to thrombin-activated FVIII. In contrast, some studies reported that FXa-activated FVIII displays less activity than thrombin-activated FVIII because of additional cleavage in the heavy chain at the site of R336. Prolonged thrombin exposure can induce non-canonical cleavages, but re-associating cleaved chains restores function. Mutational studies reveal that FXa efficiently cleaves Y346F mutants, while D361-363 mutations reduce FXa activity without affecting thrombin cleavage, highlighting distinct activation mechanisms. While the B-domain is not directly involved in FVIII activation, it may play a protective role by shielding thrombin/FXa binding and cleavage sites, thereby preventing premature activation. However, this potential function requires further investigation. [2,5,10,29,36-39]

1.4.3 FVIII - Inactivation of FVIII by Protein C: FVIIIa is rapidly inactivated due to its plasma instability, while its procofactor form is more stable, exhibiting a 100-fold lower dissociation rate. Inactivation occurs through spontaneous A2-domain dissociation or proteolytic degradation by enzymes like thrombin, FIXa, FXa, and activated protein C (aPC). aPC, a vitamin K-dependent serine protease, cleaves FVIIIa at R336 (A1-domain) and Arg562 (A2-domain), generating 54 kDa and 45 kDa fragments. FXa also inactivates FVIIIa by cleaving at K36, R336, and R1721. Binding to vWF protects FVIII from aPC. Studies show that rBDD-FVIII is more susceptible to proteolysis than full-length FVIII, correlating with increased bleeding risks. However, the reasons behind rBDD-FVIII's vulnerability remain unclear. [38,40]

1.4.4 FVIII - Clearance: FVIII is cleared from circulation post-activation, with misfolded proteins degraded via the 26S proteasome. Clearance involves receptors like the low-density lipoprotein receptor-related protein (LRP) and asialoglycoprotein receptor (ASGPR). LRP binds FVIII through A2-domain residues (484-509) and may internalize FVIII-vWF complexes. ASGPR, abundant in the liver, binds FVIII's glycosylated B-domain, influencing clearance. rBDD-FVIII shows reduced ASGPR binding, suggesting glycosylation's role in FVIII stability. [35,41-43]

1.4.5 FVIII - Interaction with vWF: FVIII circulates bound to vWF, which stabilizes it and prevents premature clearance. vWF binds FVIII's A3 and C2-domains, shielding it from proteolysis by aPC and FXa but not thrombin. Mutagenesis studies show that thrombin cleavage reduces FVIII's affinity for vWF, promoting its interaction with phospholipid surfaces during coagulation. rBDD-FVIII exhibits altered platelet-binding behavior, suggesting the B-domain influences FVIII-vWF dynamics. [33,44-50]

1.4.6 FVIII - Metal and Phospholipid Interaction

Calcium is vital for FVIII's role in coagulation, stabilizing the FVIII-vWF complex and supporting platelet adhesion. The dissociation of this complex requires >30 µM Ca^{2+}. Homology studies suggest that acidic regions in FVIII may coordinate calcium binding, influencing its structural rearrangement. Copper ions (Cu+) mediate FVIII heavy and light chain interactions, enhancing interchain affinity. Three Cu-binding sites have been identified: type 1 (A1-domain), type 3 (A3-domain), and type 2 (spanning A1 and A3), critical for structural stability. FVIIIa interacts with negatively charged phospholipids, essential for tenase complex formation with FIXa and FX. The C2-domain of the light chain mediates phospholipid binding, independent of the B-domain. This interaction aids FVIII activation, protects against inactivation, and maintains its active conformation during coagulation. [22,32,35,37]

1.5 Hemophilia A and complications

Hemophilia A is the most common severe congenital bleeding disorder, classified based on FVIII coagulant activity (FVIII: C) into severe (<1%), moderate (1-5%), and mild (5-40%) deficiency. Severe cases (50% of all) present early with bruising, joint hemorrhages, and bleeding during procedures like circumcision. Moderate (30%) and mild (20%) forms show less frequent bleeding, often triggered by trauma or surgery. Common symptoms include hemarthrosis (knees, ankles, elbows), soft-tissue hematomas, intracranial hemorrhage (ICH), and prolonged oral or post-surgical bleeding. ICH, a major risk in newborns, is more common with forceps or C-section deliveries. Bleeding episodes can occur at any age, worsening with increased physical activity. Complications include compartment syndrome, anemia, gum disease-related bleeding, and joint damage leading to arthritis. Molecularly, Hemophilia A results from mutations in the FVIII gene on the X chromosome (Xq28), encoding a 2,351-amino acid protein. This large gene contains 26 exons and 25 introns, with mutations like inversions, deletions, and missense variants causing FVIII deficiency. [51-55]

Mutations:

Hemophilia A (HA) is primarily caused by mutations in the F8 gene, with a high prevalence attributed to a significant mutation rate, especially in male germ cells. Various types of mutations have been identified in the F8 gene, including inversions, deletions, insertions, and point mutations. A comprehensive list of these mutations can be accessed through the Haemophilia A Mutation Database (HADB, also known as HAMSTeRS) at http://hadb.org.uk/. Mutation analyses indicate that the F8 gene has a mutation rate ranging from 2.5 to 4.2×10^{-5}, influenced significantly by the common intron 22 inversion and the gene’s large size. Additionally, the higher number of mitotic cell divisions during spermatogenesis contributes to an increased frequency of point mutations. Missense mutations can result in varying severities of HA and have been detected throughout the gene, from the initiation of methionine to the stop codon at position 2232. Interestingly, missense mutations within the central B-domain are rare among HA patients, suggesting that alterations in this region have minimal impact on FVIII activity. [53,56,57]

1.5.1 Hemophilia A therapies:

From plasma-derived to recombinant to Extended Half-Life (EHL) FVIII products, replacement therapies have long been the cornerstone of HA treatment. For decades, patients have relied on plasma-derived factor VIII (pd-FVIII) and recombinant factor VIII (rFVIII) products to manage their condition. Recent advancements in biotechnology have revolutionized these therapies, transitioning from standard half-life (SHL) FVIII to more advanced EHL-FVIII products, offering improved efficacy, convenience, and safety.

- **Plasma-derived Factor VIII (pd-FVIII) concentrates**

pd-FVIII concentrates have been crucial in HA treatment since the 1960s, produced through a meticulous process that begins with the collection of plasma from carefully screened donors. The plasma is pooled, fractionated to isolate FVIII, and purified using techniques like chromatography and precipitation. pd-FVIII offers several advantages, including proven efficacy, widespread availability, and cost-effectiveness, making it particularly valuable in resource-limited settings. The presence of vWF in pd-FVIII also helps stabilize FVIII in the bloodstream and may reduce the risk of inhibitor development. However, challenges remain, such as the potential for pathogen transmission, the short half-life requiring frequent infusions, and concerns about immunogenicity due to structural alterations during viral inactivation. Despite these limitations, pd-FVIII remains a critical treatment option, especially in developing countries, with products like Hemate P®, Alphanate®, and Voncento® widely used for prophylaxis, on-demand treatment, and surgical management. pd-FVIII continues to play a vital role globally, with its proven efficacy, accessibility, and cost-effectiveness ensuring its relevance in HA management, though continued innovation is essential to address its limitations and enhance patient safety. [58-62]

- **Recombinant Factor VIII (rFVIII) concentrates.**

Both rFL-FVIII and rBDD-FVIII products have transformed the treatment of HA since the 1900s, offering a safer and more reliable alternative to pd-FVIII. These products are genetically engineered to replicate the structure and function of natural FVIII using mammalian cell lines, such as Chinese hamster ovary (CHO) or baby hamster kidney (BHK) cells. The production begins with cloning the F8 gene, followed by cell culture protein production and purification to yield the final rFVIII lyophilized concentrate. rFVIII offers several advantages, but it is also shown a higher risk of inhibitor development compared to pd-FVIII, particularly in patients with severe HA, and the high cost of production. Despite these limitations, rFVIII products like Advate®, Kogenate-FS®, and ReFacto-AF® are widely used for prophylaxis, on-demand treatment, and surgical management, providing effective and reliable care for HA patients. The future of rFVIII lies in further advancements, such as EHL products using Fc fusion or PEGylation, gene therapy for potential cures, and efforts to reduce costs and improve global accessibility. [59,63]

- **Extended Half-Life Factor VIII (EHL-FVIII) Products**

The development of EHL-FVIII products represents a significant advancement in HA treatment in recent times. These therapies address the limitations of other FVIII products by enhancing pharmacokinetics, prolonging therapeutic effects, and reducing the frequency of injections. This has not only been shown to improve patient convenience but also leads to better clinical outcomes, including fewer bleeding events and improved joint health. For patients with rapid FVIII clearance, active lifestyles, or pre-existing joint damage, EHL-FVIII products offer a promising solution. EHL-FVIII products leverage two key approaches: Fc fusion technology and PEGylation. **Fc-Fusion** involves fusing the Fc region of immunoglobulin G (IgG) to the FVIII molecule, which rescues the protein from degradation and recycles it back into circulation. By exploiting this natural recycling pathway, Fc fusion extends the half-life of FVIII, enhancing its bioavailability and therapeutic efficacy. **PEGylation:** PEGylation involves attaching polyethylene glycol (PEG) molecules to the FVIII protein, which alters its physicochemical properties, reducing its immunogenicity and shielding it from clearance mechanisms. PEGylation disrupts interactions with clearance receptors in the reticuloendothelial system, thereby prolonging the protein's circulation time and improving its stability. These strategies have enabled the development of next-generation biologics with optimized pharmacokinetic and pharmacodynamic properties, offering patients longer-lasting protection with fewer injections. [59,60,63,64]

- **Gene therapy for HA**

Over the past four decades, substantial progress has been made in the therapeutic management of HA and HB. The Transformative Potential of Gene Therapy is to eliminate the need for continuous prophylactic therapy, thereby improving quality of life and restoring a sense of normalcy. The transformative impact of gene therapy research has explored various transgene delivery mechanisms, including non-viral methods such as chemical transfection, electroporation, and polymer-based techniques. However, viral vector-mediated delivery has demonstrated the most clinical promise, particularly using lentiviral vectors, adeno-associated viruses (AAVs), and gene-editing technologies. Current clinical trials and research focus on the recombinant AAV has emerged as the preferred vector for gene delivery due to its safety profile and efficiency. Nonetheless, there are three ongoing clinical trials for HA evaluating lentiviral vector-mediated gene therapy combined with autologous stem cell transplantation (NCT04418414, NCT03818763, and NCT03217032). Additionally, Valoctocogene roxaparvovec (Roctavian), developed by BioMarin, uses an AAV vector to deliver the FVIII gene and significantly reduces bleeding rates in patients, leading to FDA approval in 2023.

Giroctocogene fitelparvovec, by Pfizer and Sangamo, showed positive Phase 3 results, though Pfizer ended its partnership with Sangamo in 2024. Lentiviral gene therapy with CD34+ stem cells demonstrated sustained FVIII expression in severe HA cases. Lastly, ASC618, an AAV2/8-based therapy, is in early clinical trials with encouraging preclinical data. These studies aim to achieve durable, potentially curative treatments for HA. [59,60,62,65-67]

1.5.2 Developing inhibitors as a main challenge

An inhibitor is a high-affinity polyclonal immunoglobulin G (IgG) antibody that targets the FVIII protein. These antibodies (Abs) are predominantly of the IgG4 subclass, which does not activate the complement system. The development of FVIII inhibitors is a T-cell-dependent process involving antigen-presenting cells, B cells, and T-helper lymphocytes. FVIII antibodies can be either inhibitory or non-inhibitory. Structurally, FVIII inhibitory antibodies primarily target the A2-, A3-, and C2-domains of the FVIII molecule. When antibodies bind to these domains, they cause steric hindrance, blocking critical functional sites on FVIII. These key functional sites are essential for interactions with FIX, phospholipids, and vWF. Inhibitor Abs can simultaneously recognize multiple FVIII epitopes, and the targeted epitopes may change over time and are classified based on their kinetics and inhibitory capacity. Type I/Group A inhibitors exhibit second-order kinetics, where inhibition is dose-dependent and linear, leading to complete inactivation of FVIII. These inhibitors are more commonly associated with severe hemophilia. In contrast, Type II/Group B inhibitors show complex, nonlinear kinetics and only partially inactivate FVIII. Type II inhibitors are more frequently observed in patients with mild hemophilia or in those without congenital hemophilia who develop acquired FVIII inhibitors. [51,68,69]

Anti-A2 inhibitors can block FVIII cleavage by thrombin and disrupt FVIIIa interactions with FIXa and FX. HA patients with these inhibitors often resist immune tolerance therapy. Structurally, they remain poorly understood. Murine anti-A2 inhibitors show strong inhibitory activity targeting the highly immunogenic R484-I508 region of FVIII.

Anti-C1 inhibitors consist of different epitope regions that disrupt FVIII procoagulant activity through distinct mechanisms. Recently, the X-ray crystal structure of a “group A” anti-C1-domain inhibitor bound to the ET3i human/porcine BDD-FVIII was determined. The ET3i/2A9 Fab structure showed key hydrophobic and basic residues (F2068 and R2150) at the core interface, with the 2A9 antibody also making minor contact with the A3-domain and glycans at N1810. Notably, the complex induces a unique outward swing of the FVIII C2-domain, shifting the solvent-exposed beta-hairpin loops by ~20 Å—a distinct conformational change. [57,69-71]

Anti-C2 inhibitors revealed both “classical” and “nonclassical” epitope types. Classical epitopes inhibit FVIII procoagulant function by blocking its interaction with vWF or PS membranes. Nonclassical epitopes, on the other hand, inhibit the proteolytic activation of FVIII by thrombin or FXa, preventing its release from vWF. BO2C11, one of the most potent inhibitors, binds and sequesters the hydrophobic loops of the C2 domain, forming salt bridges with R2215 and R2220. Murine studies revealed that inhibitors with non-overlapping epitopes, such as 3E6 and G99, can act cooperatively, enhancing each other’s binding. Structural data showed that 3E6 stabilizes the opposing G99 epitope, reducing its flexibility, thus highlighting the molecular basis of cooperativity and further clarifying the membrane-binding orientation of the C2 domain. [30,69,70,72,73,73]

1.6 Integrated hybrid methodology for structural characterization

1.6.1 Building/Modeling and refining the computational models:

- **Molecular Dynamic Simulation (MD Simulation)**

MD simulations serve as a computational approach to investigate the fluctuations in atomic positions within biomolecules such as proteins or DNA over time. MD simulations bridge the gap between structure and dynamics by exploring the conformational energy landscape. Molecular simulations rely on Newtonian mechanics to model atomic motion, utilizing force fields (FF) to describe properties like bond lengths, angles, torsional movements, dihedral angles, van der Waals interactions, and electrostatic forces. These FFs play a critical role in Resolving steric clashes and optimizing rotamer geometries of the input structure during the refinement process. Since the motion of atoms within a molecule is governed by kinetic and potential forces as a function of their interactions over time, accurate FF parameterization is essential for any MD simulation. Classical MD simulations are conducted within a defined simulation cell, which represents a confined environment allowing the thermal motion of the biomolecule. This cell is equipped with specific boundary conditions (periodic boundaries or walls) and parameters like temperature and pressure. To simulate a native-like environment, the cell is typically filled with relevant solvents, such as water molecules or ions. The biomolecule of interest is simulated within a defined FF for a specific time range, which can vary from nanoseconds to microseconds. The simulation system undergoes an equilibration phase followed by the main production phase. This involves performing MD to energy-minimize the structure and bring it close to its native conformation, allowing it to differ only within the range of its natural thermal motion. The equilibration step also adapts the structure to its solvent environment and prevents the introduction of artifacts during the production phase. In this thesis, MD simulations have been extensively employed to investigate the impact of the B-domain’s glycosylations and furin-cleavage on the structural stability and integrity of FVIII protein. [74-77]

Post-MD simulation analysis:

After an MD simulation, the raw trajectory data were processed and analyzed to evaluate the structural, energetic, and dynamic aspects of the system. Structural analysis typically includes the calculation of root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and radius of gyration (Rg), which provide insights into system stability, flexibility, and compactness, respectively. Secondary structure analysis of proteins is often performed using algorithms like DSSP to track folding, unfolding, and conformational changes during the simulation. Hydrogen bond analyses are employed to investigate intermolecular interactions, solvation dynamics, and local structural arrangements. [76] A detailed description of the analysis and formulae used for the FVIII computational models in this study have been mentioned in Section 2.2.1.B.

- **Model refinement and Validation:**

Computational model refinement involves iterative improvement of an initial atomic model to achieve better agreement with experimental data while ensuring that the structure adheres to established biochemical and biophysical principles. Tools such as COOT (Crystallographic Object-Oriented Toolkit) and MiFIT play critical roles in this process. One of the most detailed aspects of model refinement involves optimizing structural geometry, which includes bond lengths, bond angles, dihedral angles, and the overall spatial arrangement of atoms. Rotamer optimization is a crucial component of refinement that Ensures side chains adopt the most favorable conformations based on their local environment. Amino acid side chains in proteins can exist in different rotameric states, defined by specific dihedral angles and higher torsional degrees of freedom. These states are constrained by steric interactions and the chemical properties of the side chains. Additionally, specific refinement challenges, such as loop building and flexible regions exhibit high flexibility. COOT provides tools for manually building or refining loops within ambiguous density regions, ensuring that the backbone geometry remains realistic and the loop fits the map appropriately. [78] Computational models refined are validated using MolProbity by Ramachandran-plot and sidechain rotamer criteria. The coordinate file of refined models is usually submitted to the MolProbity web server and checked for the number of chains, alternate conformations, and hydrogen atoms, identifying steric clashes and evaluating rotamer quality. Further, the refined models are also checked for geometry parameters such as examining bond lengths, angles, and torsion angles to ensure they adhere to ideal chemical parameters. Additionally, Ramachandran analysis is also performed to assess the φ-ψ backbone torsion angles to determine residues falling in favored, allowed, and outlier regions of the Ramachandran plot. Rotamer Analysis validates the side-chain rotamer conformations, ensuring they fit steric and energetic constraints. Identify residues with unusual rotamers that may need refinement. Clash Score analyzes steric clashes between atoms which further quantifies the number and severity of atomic overlaps. [79-81] Advanced validation metrics are mentioned further.

- **Hydrogen Bond Analysis:** Evaluate hydrogen bond geometry, including bond distances and angles, to ensure structural stability.
- **C-beta (Cβ)-Deviation Analysis:** Inspect deviations in Cβ positions from ideal geometry, which could indicate structural irregularities.
- **Validation of Metal Ions and Ligands:** If the model includes metal ions or ligands, check their coordination geometry and consistency with expected chemical properties.

Comprehensive MolProbity validation reports are generated and reviewed which provide key scores and graphical outputs for all key metrics. Key scores included

- **Ramachandran's Z-score** reflects how well the backbone dihedral angles fit expected distributions. Ideal validation metrics include Ramachandran Outliers - Less than 1 % is ideal, Rotamer Outliers - Less than 1 % is recommended for a reliable model.
- **Clash Score** is a measure of steric clashes, where atoms are too close, indicating errors in the structure. It's calculated as clashes per 1000 atoms, with a lower clash score of typically below 10 indicating high quality.
- **MolProbity Score** is a composite score combining clash score, Ramachandran statistics, and other geometric metrics. Ideal validation metrics include MolProbity Score - < 2.0 for high-quality models.

1.6.2 Biophysical Methods of Structural Characterization

- **Cryo-electron microscopy (cryo-EM)**

Cryo-electron microscopy (cryo-EM) is a cutting-edge structural biology technique that allows for the visualization of proteins and macromolecular complexes at near-atomic resolution. The scientific principles of cryo-EM include rapid vitrification, electron scattering, single-particle analysis, and computational reconstruction. The rapid vitrification of protein samples preserves their native state by flash-freezing the sample in liquid ethane (approximately -180°C). They form glass-like, amorphous ice that embeds the protein in a near-physiological environment. The vitrified sample is placed in a transmission electron microscope (TEM), where a high-energy electron beam is directed through it. Direct electron detectors (DEDs) collect electrons with higher sensitivity and reduce noise and motion blur. As electrons interact with the protein, they scatter, resulting in a two-dimensional (2D) projection image on a detector. Since

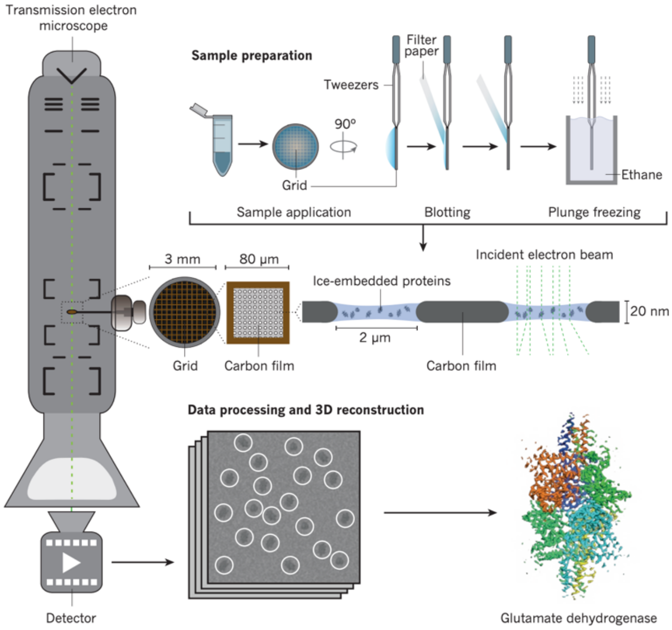


Figure 6. The fundamental principle of cryo-electron microscopy (cryo-EM)

proteins are randomly oriented in the ice, each image captures the protein from a different angle. Thousands to millions of these 2D images are gathered, with each representing a unique projection of the protein's 3D structure. The next step is computational processing to reconstruct the 3D structure, single-particle analysis (SPA). Utilizing algorithms based on principles of Fourier transforms and back-projection, these 2D images are computationally compiled to generate a 3D density map. Iterative refinement processes, such as maximum likelihood estimation and Bayesian statistics, enhance the map's resolution and accuracy. Advanced image processing software, such as RELION and cryoSPARC, has played a pivotal role in achieving near-atomic resolution (better than 3 Å) for many structures (Figure 6). This technique was specifically utilized in this thesis to resolve the structure of the FVIII protein, owing to its capability to address structural heterogeneity. [82-85] Given the heavily glycosylated and highly disordered B-domain in the full-length FVIII, methods like X-ray crystallography (XRC), which require highly ordered crystals and produce an averaged structure, would have been cumbersome. In contrast, cryo-EM's ability to capture multiple conformational states of a protein within the same sample is particularly beneficial for studying dynamic processes, such as conformational changes in large macromolecular complexes, which is especially useful for the highly disordered and flexible B-domain.

- **Atomic Force microscopy (AFM)**

Atomic force microscopy (AFM) is a powerful imaging technique widely used in protein research, enabling detailed physical measurements at the nanoscale. Initially developed for surface imaging, AFM evolved from scanning tunneling microscopy and has become a critical tool for studying protein structures and interactions. The core principle of AFM involves measuring interactions between a scanning sensor and a protein sample deposited on a flat substrate. The sensor typically consists of a silicon or silicon-nitride nanotip mounted on a micro-sized cantilever. The substrate is often a muscovite mica sheet, valued for its atomic flatness and ease of cleavage. However, alternatives such as

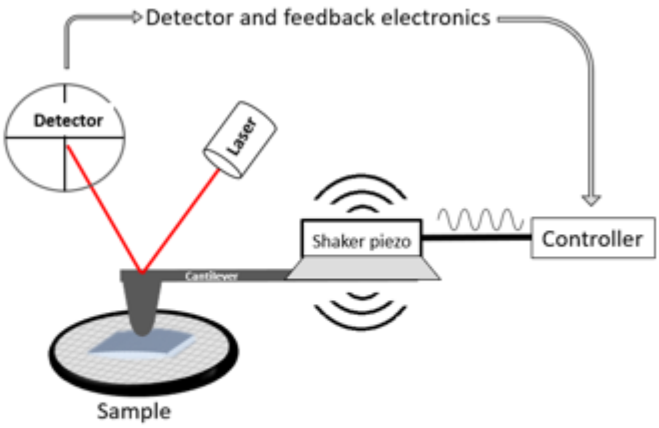


Figure 7. The basic principle of Atomic Force Microscopy (AFM)

silicon wafers, highly oriented pyrolytic graphite, and glass are also employed. These substrates can be chemically modified to facilitate the attachment of proteins, DNA, or other biomolecules, making AFM highly versatile for biological applications. For protein imaging, AFM commonly employs an optical lever detection system. In this setup, a laser beam reflects off the cantilever onto a split photodiode, enabling precise measurements of cantilever deflection. Piezoelectric tubes control the movement of the sample stage or the cantilever, allowing for surface scanning with near-angstrom resolution. An additional piezoelectric tube often operates in the vertical (z) direction, providing sub-angstrom precision for tracking height variations. A feedback loop connects the photodiode detector to the z-piezo to generate a topographic image of a protein sample, maintaining a constant interaction force between the cantilever and the sample. If the cantilever encounters a raised feature, such as a folded protein domain, the sample stage adjusts to keep the force constant, and the z-piezo's movement directly translates into a height signal (Figure 7). By scanning multiple lines, AFM constructs a detailed topographic map of the protein sample. This process is equally effective whether the z-piezo moves the sample stage or the cantilever. A key advantage of AFM for protein studies is its ability to operate in liquid environments at ambient temperatures, allowing for the imaging of proteins in near-native conditions. Sample quality is critical for high-resolution AFM imaging of proteins. Achieving a perfect AFM image of proteins is rare and depends on several factors, including the imaging mode (contact, tapping, or force-based), setup parameters (feedback setpoint, scanning rate, gains, and scan size), tip characteristics (size and durability), and sample deposition. [86-93]

1.6.3 Integrated hybrid methodology:

Integrative structural biology involves the combination of two or more analytical techniques to obtain a comprehensive and precise 3D structure of proteins or protein complexes. Accurately defining these structures is crucial for understanding protein functions and their mechanisms of action within biological systems. Foundational techniques like diffraction theory and Bragg's law, supported by Fourier transform mathematics, enabled the first protein structure determination. NMR advancements, along with innovations in cryo-EM, AFM driven by detector technology, and computational growth, have furthered structural studies. Time-resolved methods, such as spectroscopy, solution scattering, and XFEL pulses, now shed light on biomolecular dynamics. Computational tools such as MD simulation and model building, as previously mentioned, have evolved to support structural data analysis, with integrative methods combining multiple experimental approaches becoming increasingly common. As of 2024, integrative models are officially deposited in the PDB-IHM - https://pdb-ihm.org/ (previously known as PDB-Dev) database with 4-character PDB accession codes, reflecting their growing use. [94] Cryo-EM's rapid progress, especially in near-atomic resolution studies, has broadened access to complex systems like membrane proteins. Techniques like cross-linking mass spectrometry (XL-MS) and micro-electron diffraction (micro-ED) have become key for high-resolution modeling. Additionally, intrinsically disordered regions (IDRs), which represent nearly 40% of the eukaryotic proteome, are integral to cellular signaling and regulatory mechanisms, with links to numerous diseases, one such is the B-domain of the FVIII protein. Despite their critical biological roles, gaining structural and functional insights into IDRs is challenging due to their dynamic and flexible nature. This inherent disorder makes conventional structural techniques like X-ray crystallography and cryo-EM less effective. Recent progress in hybrid methodologies has opened new avenues for studying the structure of IDRs in diverse biological contexts. Hence, we utilized a similar approach of combining biophysical methods and computational analysis to explore the conformational landscape of the B-domain and full-length of the FVIII protein (Figure 8). [74,95-97]

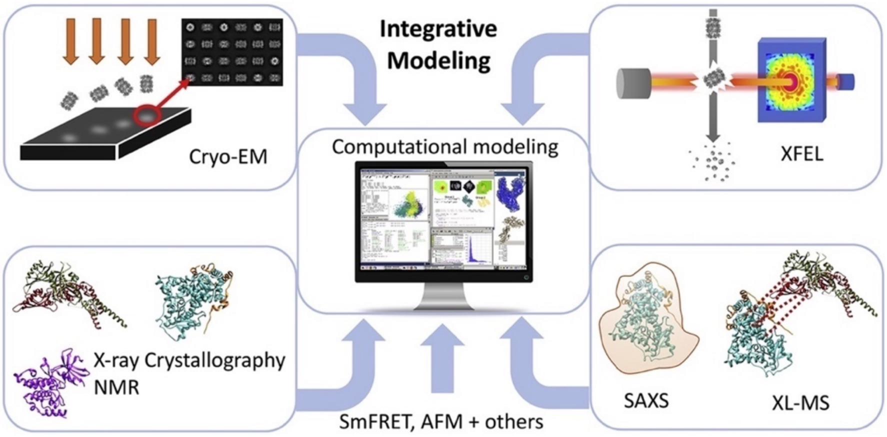


Figure 8. Overview of Integrated Hybrid Methodology

1.7 Challenges in structurally characterizing the FVIII protein:

- FVIII, being a large glycoprotein with diverse structures, is typically purified through sequential filtration and chromatography steps. Usually, the initial centrifugation and/or filtration is performed after the harvest undergoes a capture step by Size exclusion chromatography (SEC), anion exchange chromatography (AEC), or even Immunoaffinity chromatography (IAC) to remove impurities, including proteases. IAC offers high impurity removal but is costly and less stable due to the need for immobilized monoclonal antibodies. While SEC is commonly used, there is a loss of protein frequently, and the sample is usually too diluted, which needs further concentration, leading to further sample loss. Hence, pseudo-affinity purifications with protein ligands have been developed, which are cost-effective but have a moderate binding affinity, requiring multiple units and reducing yield. However, the structural diversity of FVIII protein due to heavy glycosylation and the highly flexible B-domain would complicate this process even further to obtain high yields of highly pure FVIII protein for biophysical studies. [45,98-101]

- Due to the highly disordered and flexible nature of the B-domain, capturing its full conformational landscape is challenging and may necessitate advanced structural techniques such as XRC. While cryo-EM could offer an alternative approach, it presents its own set of challenges. Cryo-EM requires a substantial amount of purified protein, which can complicate large-scale production and potentially compromise the protein's stability and quality during preparation. Moreover, rapid freezing at cryogenic temperatures might fail to preserve the full spectrum of conformational states, especially for highly dynamic regions like the B-domain, leading to incomplete structural data. [98,99,102,103]

1.8 Aims & Objectives

The B-domain is the largest structural component of plasmatic FVIII, uniquely encoded by a single large exon and exhibiting no homology to any known protein. Despite its critical role in FVIII biosynthesis, intracellular processing, trafficking, and removal of the B-domain do not appear to influence FVIII's cofactor activity. Since the B-domain is a highly disordered and heavily glycosylated domain, its structure and that of the corresponding full-length FVIII remain unknown for decades. These facts raise interest in studying the structure-functional dynamics and relevance of the B-domain. Hence, a major focus of this project is to structurally characterize both the B-domain and the full-length coagulation FVIII. This structural characterization aims to unravel the conformational dynamics of the B-domain, its interactions with other domains, and its role in FVIII biosynthesis and function. Understanding the structural basis of the B-domain could provide critical insights into its contributions to FVIII stability, secretion, and activation.

Hence, this project was designed with main aims and objects to address these unknown gaps:

- The primary objective of this project included the structural characterization of the B-domain and the full-length of FVIII using integrated techniques combining computational modeling with biophysical techniques like cryo-EM and AFM. This was performed to understand the spatial orientation of the B-domain and its degree of disorderness.

- Building upon the structure, the secondary objective focused on elucidating the structure-function relationship of the full-length protein by investigating the B-domain's role in partner interactions including vWF and Thrombin.

- Expanding the structure-functional studies, the third objective focused on investigating if and how this heavily glycosylated domain influences antibody inhibitor binding, which has been a major complication in HA.

- Additionally, the structure-function interpretation was extended to investigate the impact of the clinically reported B-domain mutation on the HA phenotype.

CHAPTER 2 - MATERIALS & METHODS

2.1 Materials

All materials, chemicals, and reagents used in this study are listed below and were of analytical grade.

2.1.1 Materials - Software and Website

SOFTWARE	WEBSITE
Protein Databases and Datasets	
UniProt	https://www.uniprot.org/uniprotkb/P00451/entry
AlphaFold2 - AF-P00451-F1-v4	https://alphafold.ebi.ac.uk/entry/P00451
RCSB-PDB	https://www.rcsb.org/
Model Archive	https://www.modelarchive.org/doi/10.5452/ma-ooxos and https://www.modelarchive.org/doi/10.5452/ma-8ueqd
Protein MD simulation, *ab initio* structure modeling, editing, and validation	
ProtParam	https://web.expasy.org/protparam/
GLYCAM-Web	https://glycam.org/
Yet Another Scientific Artificial Reality Application (YASARA) v.21.8.27.W.64	https://www.yasara.org/
PyMOL v.3.0.2	https://pymol.org/#page-top
UCSF ChimeraX v.1.8	https://www.cgl.ucsf.edu/chimerax/
Crystallographic Object-Oriented Toolkit v.1.1.10 (COOT GTK4 v.1.1.10)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Visual Molecular Dynamics (VMD) v.1.9.4	https://www.ks.uiuc.edu/Research/vmd/
MIFit v.2010.10	https://github.com/mifit
XPLOR-NIH v.3.8	https://nmr.cit.nih.gov/xplor-nih/
Sup3D	http://superpose.wishartlab.com/
APBS and pdb2pqr plugin - PyMOL v.3.0.2	https://wiki.pymol.org/index.php/Apbsplugin
MSMS	https://ccsb.scripps.edu/msms/
SpotOn	https://alcazar.science.uu.nl/cgi/services/SPOTON/spoton/
PDBePISA	https://www.ebi.ac.uk/pdbe/pisa/
GraphPad Prism v.10.3.0	https://www.graphpad.com/features
Procheck v.3.5	https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/
MolProbity	http://molprobity.biochem.duke.edu/
AFM data visualization, processing and analysis	
ScanAsyst™NanoScope v.8.10	
Gwyddion v.2.64	http://gwyddion.net/
DeStripe	http://biodev.cea.fr/destripe/
DockAFM v.1.0	http://biodev.cea.fr/dockafm/
DOT2 and AFM Assembly	
cryo-EM data visualization, processing and analysis	

Smart EPU-3	https://www.thermofisher.com/de/en/home/electron-microscopy/products/software-em-3d-vis/epu-software.html
cryoSPARC live and cryoSPARC v.3.3.2	https://cryosparc.com/
RELION v.3.1.2	https://relion.readthedocs.io/en/release-5.0/
PowerFit Server	https://alcazar.science.uu.nl/cgi/services/POWERFIT/powerfit/submit

CD data analysis

DichroWeb	http://dichroweb.cryst.bbk.ac.uk/html/home.shtml
BeStSel	https://bestsel.elte.hu/index.php

Factor VIII - Hemophilia A mutation data and analysis

EAHAD	https://f8-db-eahad.org
CDC Hemophilia-A Mutation Projects CHAMP v.2022	https://www.cdc.gov/hemophilia/mutation-project/index.html
SAAMBE-3D	http://compbio.clemson.edu/saambe_webserver/index3D.php

Plasmid designing

VectorBuilder	https://en.vectorbuilder.com/

Cell counter

NucleoView™	https://chemometec.com/software/nucleoview/

Agarose gel, SDS-PAGE electrophoresis and Western blot imaging

BioRad Image Lab v.4.1	https://www.bio-rad.com/de-de/product/image-lab-software?ID=KRE6P5E8Z

Protein purification quantification

ÄKTA Pure Unicorn™ v.7.5	https://www.cytivalifesciences.com/en/us/shop/chromatography/software/unicorn-7-p-05649?srsltid=AfmBOopUCcUV_NGhy2MmD_o1Jr743XAIiMa8T704cd4y64KA8InDK0ay
Thermo Fischer Scientific Nanodrop 1000 v.3.8.1	https://www.thermofisher.com/de/en/home/industrial/spectroscopy-elemental-isotope-analysis/molecular-spectroscopy/uv-vis-spectrophotometry/instruments/nanodrop/software.html
BioTek Gen5 microplate reader and Image software v.3.08	https://www.agilent.com/en/product/cell-analysis/cell-imaging-microscopy/cell-imaging-microscopy-software/biotek-gen5-software-for-imaging-microscopy-1623226

Mass spectrometry

ProteinLynx Global Server Data Processing	https://www.waters.com/waters/educationInstance.htm?eiid=134891878&srsltid=AfmBOoqNW8tM3O9lDxs-PPiFB-zlMgzK6MV6aQBEjGuA7dOZ-HaxYs-8

Multimer analysis

Phoresis Core - Serbia	https://www.sebia.com/softwares/software/

Dissertation preparation

Microsoft Word v.16.91	https://www.microsoft.com/de-de/microsoft-365/word?market=de
Zotero v.6.0.37	https://www.zotero.org/

2.1.2 Materials - Buffers/Reagents and Chemicals/Kits

BUFFER	RECIPE
Protein electrode running buffer	250 mM Tris base, pH 8.3 (10x), 1.92 M Glycine 1 % SDS
TAE (50x)	2 M Tris-base (pH 8.0), 1 M Glacial acetic acid 0.05 M EDTA (pH 8.0)
Tris-EDTA buffer (TE)	Invitrogen, Thermo Fisher Scientific™ Inc.
TBS (10x)	200 mM Tris base, pH 7.6, 1.379 M Sodium citrate
TBST	200 mM Tris base, pH 7.6, 1.379 M Sodium citrate, 0.1 % Tween-20
TE Buffer	10 mM Tris (pH 8.0), 1 mM EDTA
Western blocking Buffer	1x TBS, 0.1 % Tween-20 5 % Skimmed milk
Transfer buffer	25 mM Tris base, pH 8.3 , 192 mM Glycine and 20 % Methanol
1M NaCl	58.44g of NaCl in a final volume of 1 L H_2O
1M $CaCl_2$	Merck-Millipore GmbH l Sigma Aldrich Ltd.
100 mM Ammonium bicarbonate (NH_4HCO_3)	3.95 g in 500 mL of H_2O
50% acetonitrile (ACN)	Invitrogen, Thermo Fisher Scientific™ Inc.
55 mM iodoacetamide	Merck-Millipore GmbH l Sigma Aldrich Ltd. 184.96 (g/mol) in Milli-Q H_2O
REAGENT	**MANUFACTURER**
Agarose - Biozym LE	Biozym Scientific GmbH
Ampicillin sodium salt - CELLPURE®	Carl Roth®
Bovine Serum Albumin heat shock fraction, protease free, essentially globulin free, pH 7, ≥98%	Merck-Millipore GmbH l Sigma Aldrich Ltd.
LB-Broth (Lennox)	Merck-Millipore GmbH l Sigma Aldrich Ltd.
OmniPur® Ethylene Diamine Tetraacetic acid (EDTA)	Merck-Millipore GmbH l Sigma Aldrich Ltd.
Complete ULTRA Tablets, Mini, EDTA-free protease inhibitors	Merck-Millipore GmbH l Sigma Aldrich Ltd.
Dimethyl sulfoxide (DMSO), cell culture grade	PanReac AppliChem, ITW Reagents Ltd.
Precision Plus Protein Dual Color Standards	Bio-Rad Laboratories, Inc.
Methanol ≥99.9% (GC), reagent grade, ACS reagent	EMSURE®, Merck-Millipore GmbH
Midori Green Advance - Nucleic acid stain	NIPPON Genetics Europe.
TriTrack DNA Loading Dye (6X)	Thermo Fisher Scientific™ Inc.
GeneRuler 1 kb Plus DNA Ladder	Thermo Fisher Scientific™ Inc.
2X Laemmli Sample Buffer	Bio-Rad Laboratories, Inc.
Immobilon® ECL UltraPlus Western HRP Substrate	Merck-Millipore GmbH l Sigma Aldrich Ltd.
Pierce™ Coomassie Plus Bradford Reagent	Thermo Fisher Scientific™ Inc.
S.O.C. (Super Optimal broth with Catabolite repression) Medium	Invitrogen, Thermo Fisher Scientific™ Inc.
Coomassie Brilliant Blue R-250 Staining Solution	Bio-Rad Laboratories, Inc.
Glycerol reagent grade, ACS reagent	EMSURE®, Merck-Millipore GmbH
4-15% Mini-PROTEAN® TGX™ Precast Gels	Bio-Rad Laboratories, Inc.
One Shot™ TOP10 Chemically Competent E. coli	Invitrogen, Thermo Fisher Scientific™ Inc.
10x Tris/Glycine/SDS	Bio-Rad Laboratories, Inc.
HEK293T cell lines	ATCC®
Cell culture reagents	

DMEM 1X high glucose	Gibco™, Thermo Fisher Scientific™ Inc.
Lipofectamine™ 2000 Transfection Reagent	Invitrogen, Thermo Fisher Scientific™ Inc.
Fetal Bovine Serum (FBS)	Gibco™, Thermo Fisher Scientific™ Inc.
DPBS 1X, cell culture grade	Gibco™, Thermo Fisher Scientific™ Inc.
1M HEPES buffer	Biowest, The serum specialist
Opti-MEM™ I Serum Reduced Medium	Gibco™, Thermo Fisher Scientific™ Inc.
Ethanol absolute ≥99.8%, AnalaR NORMAPUR® ACS	Avantar, VWR chemicals
Penicillin-Streptomycin (10,000 U/mL)	Gibco™, Thermo Fisher Scientific™ Inc.
Trypsin-EDTA (0.25%), phenol red	Gibco™, Thermo Fisher Scientific™ Inc.
M-PER™ Mammalian Protein Extraction Reagent	Thermo Fisher Scientific™ Inc.

KIT	MANUFACTURER
Chromogenix Coamatic® Factor VIII kit	Haemochrom Diagnostica
PureLink™ HiPure Filter Midiprep kit	Invitrogen, Thermo Fisher Scientific™ Inc.
Pierce coIP kit	Invitrogen, Thermo Fisher Scientific™ Inc.
iBind™ Cards, iBind™ Solution kit	Invitrogen, Thermo Fisher Scientific™ Inc.
Hydragel-5 von-Willebrand Multimer kit	Sebia GmbH
Capturem™ IP & Co-IP kit	Takara Bio Inc.

2.1.3 Materials - Lab equipment and Labwares

LAB EQUIPMENT	MANUFACTURER
Biochemistry, Microbiology, and Molecular biology equipment	
MiniSpin ® - mini centrifuge	Eppendorf™
VF2 - Vortex shaker	JANKE & KUNKEL IKA Labortechnik
ThermoMixer™ F1.5	Eppendorf™
Precision™ General Purpose (GP) 02 Water bath	Invitrogen, Thermo Fisher Scientific™ Inc.
Witeg water bath 100°C ±0.1°C	Witeg
Centrifuge 5430 R-X	Eppendorf™
Swing Bucket Centrifuge 5804 R	Eppendorf™
MIKRO 200 centrifuge	Hettich
Refrigerated Centrifuge 3-16PK	Sigma™
C1000 Touch Thermal Cycler	Bio-Rad Laboratories, Inc.
ProBlot™ Rocker 25	Labnet International, Inc.
Nanodrop® Spectrophotometer ND 1000	Peqlab Biotechnologie GmbH
Synergy™ 2 Multi-Mode Microplate Reader	BioTek® Instruments, Inc.
Mermet UNB 100 oven bacterial incubator	Memmert
Incubation hood TH 15	Edmund Bühler GmbH
ARE Hot Plate Stirrer	Velp® Scientifica
Pioneer™ Analytical balance	OHAUS®
Kern 572-33 Precision Balance	Kern & Sohn®
Agarose gel, SDS-PAGE electrophoresis and Western blot imaging equipment	
Sub-Cell GT horizontal DNA electrophoresis cell	Bio-Rad Laboratories, Inc.
Mini-PROTEAN® Tetra cell	Bio-Rad Laboratories, Inc.

Biometra Standard Power Pack P25	Biometra Biomedizinische Analytik GmbH
PowerPac™ HC High-Current Power Supply	Bio-Rad Laboratories, Inc.
PowerPac™ Basic Power Supply	Bio-Rad Laboratories, Inc.
iBlot™ 2 Gel Transfer device	Invitrogen, Thermo Fisher Scientific™ Inc.
iBind™ Automated Western Device	Invitrogen, Thermo Fisher Scientific™ Inc.
ChemiDoc™ MP Imaging System	Bio-Rad Laboratories, Inc.
Cell culture equipment	
Heracell™ 150i CO2-Incubator	Thermo Fisher Scientific™ Inc.
LaboGene safety cabinet Scanlaf Mars 1200 Runner (Class II)	LaboGene™
Automated cell counter NucleoCounter® NC-202™	ChemoMetec A/S
Olympus Inverted Microscope CKX53 Cell Culture Microscope with Stage Drive	Evident Scientific, Inc.
Protein purification system	
ÄKTA pure™ chromatography system	Cytiva
AFM equipment	
Nanoscope V Multimode 8	Bruker
DriveAFM	Nanosurf
cryo-EM equipment	
Tecnai 12 LaB6 electron microscope	Thermo Fisher Scientific™ Inc.
Talos™ L120C TEM	Thermo Fisher Scientific™ Inc.
Vitrobot mark IV plunge freezer	Thermo Fisher Scientific™ Inc.
Krios™G4 Cryo-TEM	Thermo Fisher Scientific™ Inc.
Glacios™ 2 Cryo-TEM	Thermo Fisher Scientific™ Inc.
CD spectrometry	
JASCO J-715 CD spectrophotometer	Jasco Deutschland GmbH
Coagulation and Multimer analysers	
CN-6000 coagulation analyser	Siemens Healthineers

LABWARE	MANUFACTURER
Biochemistry, Microbiology, and Molecular biology	
Peha-soft® nitrile powder-free	HARTMANN AG
microPipettes and Pipette Tips	Eppendorf™
Centrifuge Tubes (15 mL, 50 mL)	Corning®
Microcentrifuge Tubes (1.5 mL, 2 mL)	Eppendorf™
Petri Dishes (100 mm, 150 mm)	Corning®
TC-Platte 6 well, Standard	SARSTEDT AG & Co. KG
Cellstar® 96-Well cell culture microplates	Greiner Bio-One
TC- Flasks T25, T75 Standard	SARSTEDT AG & Co. KG
Serological pipettes ROTILABO® Sterile, 5 mL, 10 mL, 25 mL	Carl Roth®
MicroAmp™ Optical 8-Cap Strips	Applied Biosystems™, Thermo Fisher Scientific™ Inc.
MicroAmp™ Optical 8-Tube Strip, 0.2 mL	Applied Biosystems™, Thermo Fisher Scientific™ Inc.
Via2-Cassette™	ChemoMetec A/S

Agarose gel, SDS-PAGE electrophoresis and Western blot imaging equipment	
iBlot™ 2 Transfer Stacks, PVDF, mini	Invitrogen, Thermo Fisher Scientific™ Inc.
iBind™ Cards	Invitrogen, Thermo Fisher Scientific™ Inc.
Amersham™ Hybond® P Membrane for Western Blot, PVDF	Merck-MilliporeSigma Ltd.
Protein purification	
Superdex™ 200 Increase 10/300 GL	Cytiva
Superose™ 6 Increase 10/300 GL	Cytiva
Anion exchange chromatography columns, HiTrap™ Q FF	Cytiva
Ultrafree ® -MC centrifuge filter units with microporous membrane	Merck-MilliporeSigma Ltd.
Vivaspin® 500 Centrifugal Concentrators - 50 kDa	Sartorius Stedim Lab Ltd.
Amicon® Ultra Zentrifugenfilter, 50 kDa MWCO	Merck-MilliporeSigma Ltd.
BD Plastipak™ fine dosing 1 mL syringes	BD - Becton, Dickinson and Co.
Atomic Force Microscopy	
Triangular ScanAsyst-Fluid probes	BrukerAFMProbes, Camarillo
ScanAsyst-Air-HR probes	BrukerAFMProbes, Camarillo
WM0.6 AuD probes	Nanosurf AG
cryo-EM	
Quantifoil® R1.2/1.3 200 Mesh, Cu	Quantifoil®
Quantifoil® R1.2/1.3 Cu/Rh 200	Quantifoil®
CD spectrometry	
0.1 cm UltraMicro-Küvette	Hellma®

2.1.4 Materials - Computational models generated

MODEL NAMES	MODEL DESCRIPTION
AF-FL-FVIII	AlphaFold full-length Factor VIII model
Gly-FL-FVIII	Glycosylated full length Factor VIII model
NGly-FL-FVIII	Non-glycosylated full length Factor VIII model
Gly-BDD-FVIII	Glycosylated B-domain deleted Factor VIII model
Gly/FC-FL-FVIII	Glycosylated furin cleaved (R1667) full length Factor VIII model
NGly/FC-FL-FVIII	Non-glycosylated furin cleaved (R1667) full length Factor VIII model
Gly/FC-FL-FVIII-D’-D3 complex	Glycosylated furin cleaved (R1667 and R1332) full length Factor VIII with vWF-D’ and D3 domains complex model
Gly/FC-FL-FVIII	Final simulated, refined and validated - Glycosylated furin cleaved (R1667) full length Factor VIII model
Gly-BDD-FVIII	Final simulated, refined and validated - Glycosylated B-domain deleted Factor VIII model

Gly/FC-FL-FVIII-D’-D3 complex	Final simulated, refined and validated - Glycosylated furin cleaved full length Factor VIII with vWF-D’ and D3 domains complex model

FVIII models deposited in the model archive

Gly/FC-FL-FVIII **-** Model Archive ID **- 10.5452/ma-ooxos**
Gly/FC-FL-FVIII-D’-D3 complex **-** Model Archive ID **- 10.5452/ma-8ueqd**

2.2 METHODS - Computational studies

2.2.1 Computational studies - Models of FL-FVIII, BDD-FVIII and FL-FVIII-vWF complex structures

2.2.1.A Model building FL-FVIII, BDD-FVIII, and FL-FVIII-vWF complex computational models

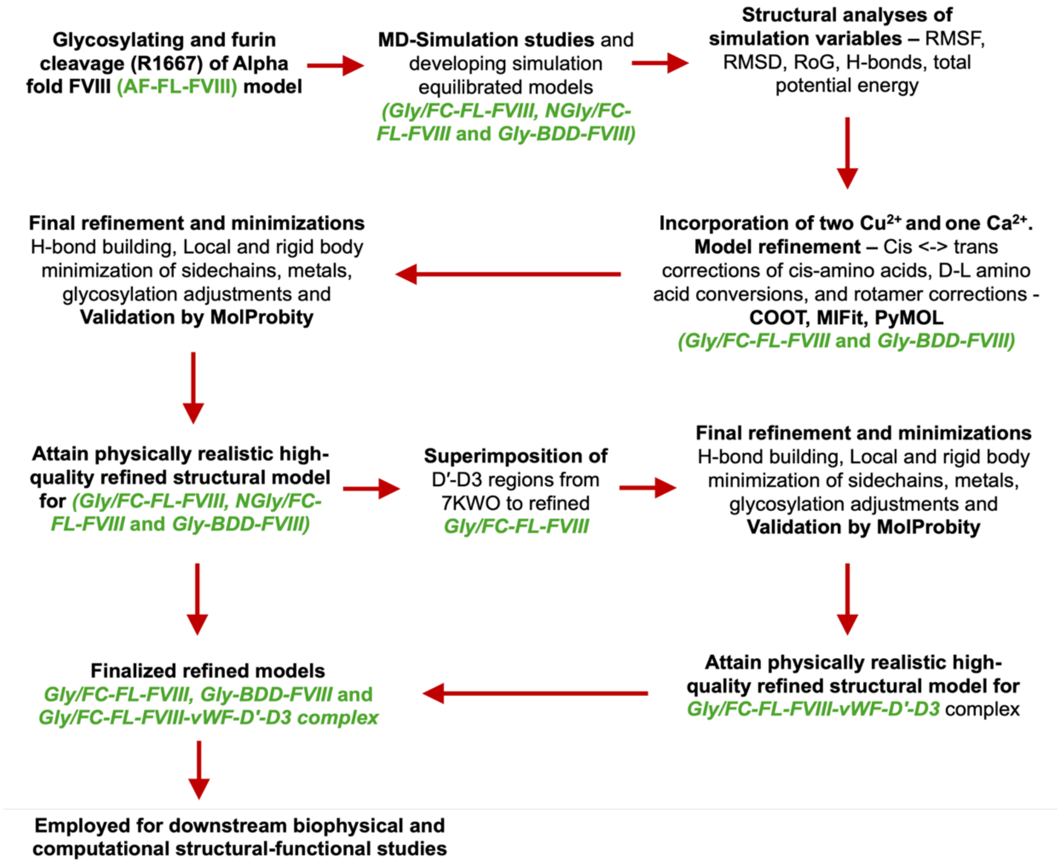


Figure 9. Stepwise methodology of model building and refinement of computational models. The workflow of developing *Gly/FC-FL-FVIII, Gly-BDD-FVIII, and Gly/FC-FL-FVIII-vWF-D'-D3* computational models involved glycosylation and furin cleavage of the AlphaFold FVIII model, followed by MD simulations, structural refinements, and validations. Final and refined *Gly/FC-FL-FVIII, Gly-BDD-FVIII, and Gly/FC-FL-FVIII-vWF-D'-D3* models generated were utilized for downstream biophysical and computational analyses.

- **AlphaFold of full-length factor VIII (AF-FL-FVIII):** The FL-FVIII computational model was built upon the model extracted from the AF2 database (AF-FL-FVIII) as a primary reference. AlphaFold2 was employed to predict the FL-FVIII protein structure based on its amino acid sequence and data from the UniProt database

(Accession ID: P00451), ensuring accuracy in the process. The AlphaFold2 framework integrates multiple databases, such as UniRef90, MGnify, and PDB70, to provide sequence alignments and structural templates. The structural modeling process in AlphaFold2 involves iterative refinement of the 3D structure using its neural network, which processes both multiple sequence alignments (MSA) features and template alignments. Independent structural models generated by AF2 are usually deposited in the AF2 database and the model is scored for confidence using the predicted Local Distance Difference Test (pLDDT). The pLDDT score serves as a quantitative measure of prediction reliability, with higher scores indicating greater accuracy. The predicted AF-FL-FVIII model was represented through the color-coded model to predict confidence using pLDDT scores to indicate prediction reliability: blue (>90) for very high confidence, cyan (70-90) for high, yellow (50-70) for moderate, and orange (<50) for low. Additionally, a Predicted Aligned Error (PAE) graph was plotted which evaluates confidence in the relative positions of residues, highlighting inter-domain flexibility and alignment accuracy. [104-107]

- **Glycosylations:** B-domain is a heavily glycosylated domain with both N- and O-linked glycosylations. N-linked glycosylations target asparagine (N) residues within the N-X-S/T sequence motif, where X represents any amino acid except proline, while O-linked glycosylations target serine (S) or threonine (T) residues. To represent the native physiological glycosylation state and to further understand the role of glycosylation in maintaining the structure of FVIII, N-linked glycosylations were incorporated into the AF-FL-FVIII model (Gly-FL-FVIII) using a GLYCAM webserver. GLYCAM-WEB is dedicated to simplifying the prediction of the three-dimensional (3D) structure of carbohydrates and macromolecular structures involving carbohydrates. [108-110] It provides tools to facilitate molecular modeling of oligosaccharides and larger molecular structures containing glycans. Valuable information regarding the type and sites of glycosylations commonly observed in FVIII protein was obtained from the previous literature studies. [111-113] Although FVIII contains 24-N-linked glycosylation sites and 5 O-linked glycosylation sites, it is found that 19 out of these 24 putative sites are occupied with high-mannose and hybrid glycan structures, and not enough data about the O-linked glycosylation is present. Glycam's library was utilized to select individual sites and pre-designed glycan templates - DNeu5Aca2-6DGalpb1-4DGlcpNAcb1-6[DNeu5Aca2-6DGalpb1-4DGlcpNAcb1-2]DManpa1-6[DNeu5Aca2-6DGalpb1-4DGlcpNAcb1-2[DNeu5Aca2-6DGalpb1-4DGlcpNAcb1-4]DManpa1-3]DManpb1-4DGlcpNAcb1-4[LFucpa1-6]DGlcpNAcb1-OH. The chosen N-linked glycans, create a β-1,4 bond between the N-acetylglucosamine (GlcNAc) base and the nitrogen atom of the (Asn-N) residue. Glycam FF - GLYCAM06, a specialized FF designed for modeling carbohydrates and glycoproteins, and its respective interactions in molecular simulations was assigned to the FVIII-glycoprotein complex. Further, energy minimization, embedded in GLYCAM-WEB was also performed on the FVIII-glycoprotein complex to resolve any steric clashes and stabilize the molecule, which was further prepared for downstream modifications and MD simulation studies. [108,109] Additionally, the initial AF-FL-FVIII model without glycosylation (NGly-FL-FVIII) was analyzed comparatively to investigate the role of glycosylation in FVIII.

- **Furin cleavage:** Furin is a pro-protein convertase that cleaves FVIII at specific sites, converting the inactive pro-FVIII to a mature FVIII. Furin cleavage is known to occur at several sites, but primary PACE/Furin cleavage occurs at R1667 followed by secondary yet prominent furin cleavage at R1332 leading to the formation of different lengths of B-domain and $HC_{(A1\text{-}A2\text{-domains})}$. [114-118] To further mimic these native physiological conditions in our models, furin cleavage at the R1667 was performed using YASARA 21.8.27.W.64 (Yet Another Scientific Artificial Reality Application-21.8.27.W.64). [119,120] The cleavage was performed between R1667 and E1668 by breaking the peptide bond, thus disrupting the shared electrons between the two residues leading to hydrolysis. This energy-driven process formed two distinct peptide fragments with new N-termini containing $HC_{(A1\text{-}A2\text{-}B\text{-domains})}$ and C-termini containing $LC_{(A3\text{-}C1\text{-}C2\text{-domains})}$ thus triggering a conformational change. Following the cleavage, the resulting structure was subjected to energy minimization (EM) to remove any steric clashes or unfavorable interactions caused by the bond breakage. This ensured that the cleaved FVIII structure was energetically stable and suitable for further analysis. Both the glycosylated (Gly-FL-FVIII) and non-glycosylated (NGly-FL-FVIII) models mentioned earlier, were subjected to the furin cleavage resulting in forming glycosylated-furin cleaved (Gly/FC-FL-FVIII) and non-glycosylated-furin cleaved (NGly/FC-FL-FVIII) models (MD simulation and EM of the models further explained in Section 2.2.1.B).

- **Generating B-domain deleted-FVIII model:** BDD-FVIII is one of the most predominant, prominent forms of FVIII specifically necessary for the cofactor activity of the FVIII. Even though several crystal and cryo-EM structures corresponding to the BDD-FVIII have been deposited in the RCSB-PDB database, most of them lack certain residues in between. [26,30,32-34,121] Hence glycosylated BDD-FVIII (Gly-BDD-FVIII) model was generated by deleting the residues from 759 to 1667 corresponding to the B-domain from the Gly/FC-FL-FVIII model (as detailed previously) using YASARA. The prominent 6 glycosylation sites, N60, N258, N601, N1704, N1829, N2137, containing N-glycans present in A1-,A2-, A3- and C1-domains were retained. Subsequently, MD simulation and EM were performed to relax the structure and relieve any steric clashes or unfavorable interactions that might have occurred due to the modification to generate the Gly-BDD-FVIII model (MD simulation and EM of the models further explained in Section 2.2.1.B).

- **Generating FL-FVIII-vWF complex model:** vWF a chaperon protein, complexes with FVIII as a non-covalent complex is found in circulation and also stabilizes and protects FVIII from premature degradation. Although a recent cryo-EM structure BDD-FVIII-vWF complex containing D'-D3 domains attached to the C1-C2-domains of FVIII-XTEN Fc fusion protein was characterized, (PDB-ID: 7KWO) [33] a FL-FVIII-vWF complex structure containing the B-domain is still missing. To overcome this, we superimposed D'-D3 domains from the 7KWO structure onto the simulated Gly/FC-FL-FVIII structural model (MD simulation and EM of the models further explained in Section 2.2.1.B) using the SUP2 algorithm. [122] The D'-D3 domains were primarily aligned with the corresponding regions in the Gly/FC-FL-FVIII to generate the Gly/FC-FVIII-vWF-D′-D3 complex model with primary furin cleavage at R1667. The SUP2 algorithm works by aligning the secondary structure elements (i.e., alpha-helices, beta-strands, loops), backbone atoms, and side-chain orientations of the D'-D3 domains from 7KWO and identifying the optimal spatial alignment. The algorithm optimizes the orientation of the domains to minimize the RMSD between the structures. After performing the superimposition, the RMSD of the superimposed structure was estimated to determine the quality of the alignment. [123] A low RMSD value (typically under 2 Å) indicates a good fit and suggests that the superimposition is accurate, and further EM was performed to relax the structure and relieve any steric clashes or unfavorable interactions. The superimposed structure was then also utilized to perform primary and secondary furin cleavage at the R1667 and R1332 to generate Gly/FC-FL-FVIII-vWF-D′-D3 and understand the role of furin cleavage (further explained in Section 2.2.1.B).

2.2.1.B Molecular Dynamic simulation and energy minimization of the FL-FVII, BDD-FVIII, and FL-FVIII-vWF complex models

All three models - Gly/FC-FL-FVIII, NGly/FC-FL-FVIII, and Gly-BDD-FVIII were subjected to MD simulations using the YASARA structure suite. First, a simulation cell large enough to encompass the entire protein, including a buffer region for solvent was set. Typically, 10-12 Å of space between the protein and the edge of the box is recommended to avoid interactions with periodic images. Add Water feature was used to solvate the protein models. Add ions feature can add counter-ions to neutralize the total charge of the system. Add Ions and specify the desired ion concentration (e.g., 0.15 Molar (M) NaCl) to simulate physiological conditions. The position of atomic coordinates was refined through 500 picoseconds (ps) with YAMBER3 FF parameters in YASARA to eliminate steric clashes and optimize rotamer geometry, yielding energy-minimized structures. The resulting lowest energy structures were selected for further simulations utilizing the md_run macro embedded in YASARA. Briefly, a simulation cell with periodic boundaries and a minimum distance of 20 Å from protein atoms was employed with explicit solvent. The AMBER03 FF was used within an NPT ensemble, incorporating a long-range particle-mesh Ewald (PME) potential and a cutoff distance of 7.86 Å. Hydrogen bond networks were optimized according to the method of Hooft et al. [124] The simulation cell was filled with water at a density of 0.997 gram/milliliter (g/mL), maintaining a maximum sum of all bumps per water molecule at 1.0 Å. The net charge of the simulation cell was neutralized with a final NaCl concentration of 0.9 % (weight/volume). The entire system underwent energy minimization by steepest descent to alleviate conformational stress, followed by simulated annealing minimization until convergence. MD simulations were conducted at 298 K to observe the effects of glycosylation and furin cleavage, with simulation durations of >100 nanoseconds (ns) for the production phase for all three Gly/FC-FL-FVIII, NGly/FC-FL-FVIII and Gly-BDD-FVIII models. [119,120] This yielded simulation equilibrated *Gly/FC-FL-FVIII and NGly/FC-FL-FVIII, Gly-BDD-FVIII* models. These 3 models were then subjected to post-simulation analysis as detailed below structural refinement as mentioned in Section 2.2.1.C and validations as discussed in Section 2.2.1.D.

Post-MD simulation analysis

Post-simulation analysis was performed on the simulated models using the md_analyse in the YASARA tool to compare the three *Gly/FC-FL-FVIII, NGly/FC-FL-FVIII, and Gly-BDD-FVIII* models. The trajectory files capturing snapshots of the protein structure at specified intervals (every 100 ps) were utilized for post-simulation analysis. Total potential, kinetic energy, and temperature data were used to assess the stability of the system. Subsequent comparative structural analysis utilizing RMSD, RMSF, Rg, secondary structure changes, and H-bonds including protein-ligand interactions, the Total potential energy of the system throughout the simulation was analyzed to assess conformational stability and dynamics. The following parameter with their respective formulae is explained in detail below.

- **Root Mean Square Deviation (RMSD):** RMSD was calculated by superimposing the final simulation trajectory *Gly/FC-FL-FVIII, NGly/FC-FL-FVIII, and Gly-BDD-FVIII* structure onto the initial simulation trajectory through least-squares fitting, aligning selected atoms (e.g., backbone C, C-alpha, N). Later the average distance between corresponding atoms after alignment was measured. Lower RMSD values indicate higher structural similarity, providing a quantitative measure of deviation between structures.
  The RMSD formula:

  $$RMSD = \sqrt{\frac{\sum_{i=1}^{n} R_i * R_i}{n}}$$

  Ri is the vector linking the positions of atom i [of N atoms] in the reference snapshot and the final snapshot after optimal superposition.

- **Root Mean Square Fluctuation (RMSF):** RMSF was quantified by superimposing the final simulation trajectory *Gly/FC-FL-FVIII, NGly/FC-FL-FVIII, and Gly-BDD-FVIII* structure onto the initial simulation trajectory the flexibility of individual atoms or residues over a simulation trajectory. It was calculated as the average deviation of

an atom's position. Higher RMSF values indicate greater mobility, useful for identifying flexible or stable regions in a structure.
The RMSF formula :

$$RMSF_i = \sqrt{\sum_{j=1}^{3}\left(\frac{1}{N}\sum_{k=1}^{N} P_{ikj}^2 - \overline{P_{ij}}^2\right)}$$

RMSF of atom i with j running from 1 to 3 for the x, y, and z coordinate of the atom position vector P and k running over the set of N-evaluated snapshots.

- **Radius of Gyration (Rg):** The radius of gyration describes the compactness of a molecule and is calculated as the mass-weighted root mean square distance of atoms from their center of mass.
  The Rg formula :

$$Radius_{gyr,Mass} = \sqrt{\frac{\Sigma_{i=1}^{N} Mass_i(\vec{R}_i - \vec{C})^2}{\Sigma_{i=1}^{N} Mass_i}}$$

  C is the center of mass, and Ri is the position of atom i of N.

- **Number of Hydrogen bonds (H-bonds):** A hydrogen bond is considered stable if the donor-acceptor distance is ≤ 3.5 Å, and the donor-H-acceptor angle is ≥ 120°. The hydrogen bonds (H-bonds) are calculated using geometrical criteria to ensure the detection of significant and geometrically favorable H-bonds in the structure. During the analysis, one H-bond per hydrogen atom is assigned at most. The following formula yields the bond energy in [kJ/mol] as a function of the Hydrogen-Acceptor distance in [A] and scaling factors:

$$Energy_{HBo} = 25 * \frac{2.6 - max(Dis_{H-A}, 2.1)}{0.5} * Scale_{D-H-A} * Scale_{H-A-X}$$

$$Scale_{D-H\ A} = \begin{cases} 0 & \text{in range 0..100 degrees} \\ 0..1 & \text{in range 100..165 degrees} \\ 1 & \text{in range 165..180 degrees} \end{cases} \qquad Scale_{H\ A-X} = \begin{cases} 0 & \text{in range 0..85 degrees} \\ 0..1 & \text{in range 85..95 degrees} \\ 1 & \text{in range 95..180 degrees} \end{cases}$$

$$Scale_{H\ A-H} = \begin{cases} 0 & \text{in range 0..75 degrees} \\ 0..1 & \text{in range 75..85 degrees} \\ 1 & \text{in range 85..180 degrees} \end{cases}$$

  The first scaling factor depends on the angle formed by the Donor-Hydrogen-Acceptor, second scaling factor is derived from the angle formed by Hydrogen-Acceptor-X, where X is the atom covalently bound to the acceptor (if X is a heavy atom or is a hydrogen)

- **Total potential energy (TPE):** The total potential energy of a protein structure is calculated by summing the various energy components that account for different interactions within the system. These components included van der Waals

interactions, electrostatic forces, bond stretching and bending, torsional energy, and solvation effects. The total energy is minimized to optimize the protein's conformation.
The TPE formula

$$E_{\text{total}} = E_{\text{vdW}} + E_{\text{elec}} + E_{\text{bond}} + E_{\text{torsion}} + E_{\text{solvation}}$$

E_{vdW} - van der Waals energy, E_{elec} - electrostatic energy, E_{bond} - the bond energy, $E_{torsion}$ - the torsional energy, $E_{solvation}$ - the solvation energy.

- **Secondary structure content (SSC):** The secondary structure of proteins in YASARA is predicted by analyzing backbone dihedral angles (phi/psi) and spatial interactions between residues. The software typically utilizes the DSSP (Define Secondary Structure of Proteins) algorithm, which categorizes regions of the protein into α-helices, β-sheets, or coils based on these parameters. No specific formula is applied; instead, the secondary structure is assigned based on the dihedral angles and residue interactions, typically using the DSSP algorithm.

2.2.1.C Model Refinement of the FL-FVII, BDD-FVIII, and FL-FVIII-vWF complex models

After the initial simulation, refinement steps were performed to optimize the structure and ensure accurate metal coordination within the protein and to build high-quality structural models of *Gly/FC-FL-FVIII and Gly/FC-FVIII-vWF-D'-D3* models. This process included the addition of the metal ions in the protein structure based on predefined binding sites and reference coordinates from BDD-FVIII crystal structures (PDB-ID: 3CDZ and 2R7E). The models with metals were further energy minimized and refined to ensure the structural models are geometrically accurate and reliable for further analysis.

- **Superimposition of Cu^{2+} and Ca^{2+} to the *Gly/FC-FL-FVIII* model:** To model the metal-binding sites and perform energy minimization on FVIII, several steps were followed. First, reference files for copper (Cu^{2+}, Cu^{2+}) and calcium (Ca^{2+}) binding sites from 2R7E and 3CDZ, were prepared. Using the *sup* the metal ions were positioned onto *Gly/FC-FL-FVIII* and further refined in the structure. The metal coordination and its refinement onto the superimposed models were performed by adding patches for disulfide bridges and Histidines, followed by running the XPLOR-NIH v.3.8 energy minimization.
- **Refinement of the models:** As part of the model building and refinement process of *Gly/FC-FL-FVIII and Gly/FC-FVIII-vWF-D'-D3 complex* models, local and side chain adjustments, minimizations were performed using COOT, MIFit, PyMOL, and XPLOR-NIH v.3.8 to enhance and achieve a stable, accurate, realistic model with high accuracy and quality. [78] COOT and MIFit were utilized for Cis-trans corrections

of cis-amino acids, D-L amino acid conversions, rotamer corrections, manual optimization to resolve steric clashes and improve side chain geometry, and H-bond building, metal optimization, rigid body and local energy minimization, sidechain minimization and optimizing overall structure energetics was performed using XPLOR-NIH v.3.8 and PyMOL. [125-127] Post-refinement and remodeling, the models were analyzed and validated with MolProbity, as mentioned below in Section 2.2.1.D. These refinements contributed to building high-quality models with accurately placed metal ions, ready for further analysis.

2.2.1.D Model validation of the FL-FVII, BDD-FVIII, and FL-FVIII-vWF complex models

The refined structural models *Gly/FC-FL-FVIII,* and *Gly/FL-FVIII-vWF-D'-D3* complex models were validated MolProbity. MolProbity is a comprehensive structural validation tool that evaluates the quality of protein structures, providing insights into stereochemistry, geometry, and atomic contacts. All the parameters mentioned earlier in Section 1.4.1 were considered for the validation to provide a comprehensive report. Further analysis and interpretation of the scores were performed by comparing the validation report for *Gly/FC-FL-FVIII, Gly-BDD-FVIII, and Gly/FL-FVIII-vWF-D'-D3* with AF-FL-FVIII. Outliers in the Ramachandran plot, high steric clash scores, or unusual rotamer conformations were addressed by revisiting the modeling process or refining the structure using energy minimization tools. [79-81]

2.3 Methods - FVIII protein biochemistry studies

2.3.1 FVIII protein biochemistry studies - Characterisation of FL-FVIII, BDD-FVIII, and FL-FVIII-vWF complex proteins

2.3.1.A Commercial recombinant and plasma-derived FVIII products

Recombinant full-length FVIII (rFL-FVIII) protein was obtained from commercially available product - Kovaltry^{TM}, Bayer Inc. Plasma-derived full-length FVIII (pdFL-FVIII) protein was obtained from commercially available product - Haemoctin^{TM}, Biotest Pharma GmBH. Recombinant B-domain deleted FVIII (rBDD-FVIII) protein was obtained from commercially available product - Nuwiq^{TM}, Octapharma GmbH. [63,128,129]

FVIII variant	FVIII products	Manufacturer	Half-life in hours (h)	Source of production	Stabilizer
Recombinant full-length FVIII (rFL-FVIII)	Kovaltry^{TM}	Bayer Inc	13.4	Baby hamster kidney cell (BHK)	Sucrose/ Albumin
Recombinant B-domain deleted FVIII (rBDD-FVIII)	Nuwiq^{TM}	Octapharma GmbH	17.1	Human Embryonic Kidney 293 (HEK293)	Sucrose/ Arginine
Plasma derived full-length FVIII (pdFL-FVIII)	Hemoctin^{TM}	Biotest Pharma GmBH	12.2-17.9	Human plasma donations and pooled human plasma	-

Table 1. Overview of FVIII commercial products utilized for the study. Commercially available FVIII products, including recombinant full-length FVIII (Kovaltry), recombinant B-domain deleted FVIII (Nuwiq), and plasma-derived full-length FVIII (Hemoctin) and their details, including manufacturers, half-lives, sources of production, stabilizers are included.

- **Kovaltry^{TM} - rFL-FVIII:** Manufactured by Bayer Inc, is a recombinant unmodified full-length FVIII (rFL-FVIII) product produced in genetically engineered baby hamster kidney (BHK) cells with Sucrose/Albumin as a stabilizer purified via Imuno-affinity Chromatography (IAC)/ Ion-Exchange Chromatography (IEX) /Solvent/Detergent treatment (SD) /Ultrafiltration (UF) with half-life of 13.4 hours (h).
- **Nuwiq^{TM} - rBDD-FVIII:** Developed by Octapharma, is a recombinant B-domain deleted FVIII (rBDD-FVIII) derived from a human cell line (HEK293) with Sucrose/Arginine as a stabilizer and it is further purified via IAC/IEC/SD/NF/SE with a half-life of 17.1 h.
- **Haemoctin^{TM} - pdFL-FVIII:** Produced by Biotest Pharmaceuticals, is a plasma-derived FVIII concentrate obtained from screened human plasma donations and pooled human plasma purified through human plasma through IAC/IEC/SD/UF with a half-life of 12.2-17.9 h.

2.3.2 FVIII protein biochemistry studies - FVIII protein purification strategies

2.3.2.A FVIII protein purification strategies - Size exclusion chromatography (SEC)

Size Exclusion Chromatography involves the separation and purification of proteins by their molecular weight/size. A size exclusion column contains a porous matrix with various pore sizes allowing larger proteins to elute earlier as they are unable to enter smaller pores and smaller proteins as they enter smaller pores inside the matrix, take less time to pass through the column and elute later. An overview of the FVIII SEC purification strategy is depicted in Figure 10.

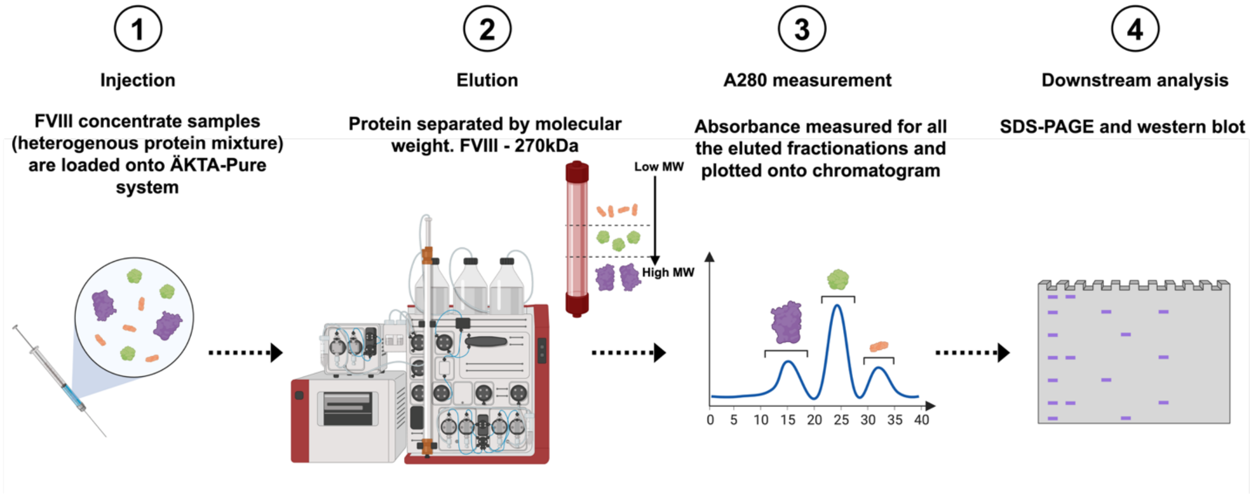


Figure 10. SEC purification of FVIII protein. Schematic representation of size exclusion chromatography (SEC) performed on an ÄKTA Pure system. SEC purification of FVIII protein was performed to obtain high pure FVIII protein from commercially available products based on its

molecular weight using Superdex™ 200 Increase 10/300 GL and Superose™ 6 Increase 10/300 GL columns. The obtained fractionations were further verified by downstream analysis.

Purification of rFL-FVIII and rBDD-FVIII was performed using Superdex™ 200 Increase 10/300 GL (Cytiva) and pdFL-FVIII was purified using Superose™ 6 Increase 10/300 GL (Cytiva) connected to an FPLC of ÄKTA pure™ (Cytiva). The columns were equilibrated with at least one column volume (CV) of filtered and degassed running buffer 1X PBS. Degassing of 1X PBS for purification was performed to remove dissolved gases, such as oxygen and carbon dioxide, which could interfere with protein stability and chromatographic performance. This was achieved by placing the buffer under a vacuum to gently sparge the solution and further utilize it for purification purposes. 500 µL of the reconstituted crude rFL-FVIII, rBDD-FVIII, and pdFL-FVIII protein samples were applied by injecting them into a 500 µL loop connected to the ÄKTA pure system. The flow-through (FT) was collected by an automated fractionator in fraction volumes of 1 mL. Fractions corresponding to the vials containing the FVIII and FVIII-vWF complex proteins in rFL-FVIII, rBDD-FVIII, and pdFL-FVIII runs were concentrated as described in Section 2.3.3. A and Section 2.3.3.B were then used for further analysis. For the calculation of molecular weight estimates from analytical SEC elution profiles, the following formula was applied to calculate the K_{av} value*:* $V_e - V_o/V_t$ - V_o Here, V_e, V_o, and V_t denote the elution volume of the analyte, the void volume and the total volume of the column, respectively. K_{av} is the partition coefficient and can be described as the proportion of pores available to the analyte. Since there is a linear relationship between the K_{av} of a molecule and the logarithm of its molecular weight, a calibration curve was also generated by running a molecular weight standard. Additionally, small amounts of each fraction were analyzed by SDS-PAGE and Coomassie staining as mentioned in Section 2.3.3.C, and molecular weight was determined here as well. The purified fractions were also subjected to mass-spectrometry analysis for their purity as mentioned in Section 2.3.3.E.

2.3.2.B FVIII protein purification strategies - Co-Immunoprecipitation (co-IP) and Ion-Exchange chromatography (IEX)

- **Co - Immunoprecipitation (co-IP) of rFL-FVIII, pdFL-FVIII:** Co-IP of FVIII protein was performed using Pierce™ Co-Immunoprecipitation Kit (Invitrogen, Thermo Fischer Scientific). 75 µg of FVIII antibodies (Ab) - PAHFVIII-S FVIII Ab (Haematologic Technologies Prolytix and CellSystems®) and Sheep Anti-Human Factor VIII, affinity purified IgG - SAF8C-AP FVIII Ab (Affinity™ Biologicals Inc.) were incubated and immobilized on AminoLink Plus Resin, using coupling buffer and incubated for 1-2 h at RT to allow Ab binding. The Ab-immobilized resin was further washed and neutralized to make sure of the efficient Ab binding. rFL-FVIII and rBDD-FVIII proteins were then added and incubated in these Ab-immobilized resins. After incubation, the samples were centrifuged at 2,000 x g for 5 minutes at 4°C to pellet the resins. The resin-Ab-protein complex was then washed to remove non-specifically bound proteins. Finally, the FVIII bound proteins were then eluted and the eluted fractions were analyzed by SDS-PAGE with Coomassie staining to assess protein purity and Western blot analysis was performed to confirm the identity of the purified FVIII protein as mentioned in Sections 2.3.3.C, 2.3.3.D and 2.3.3.E (Figure 11).

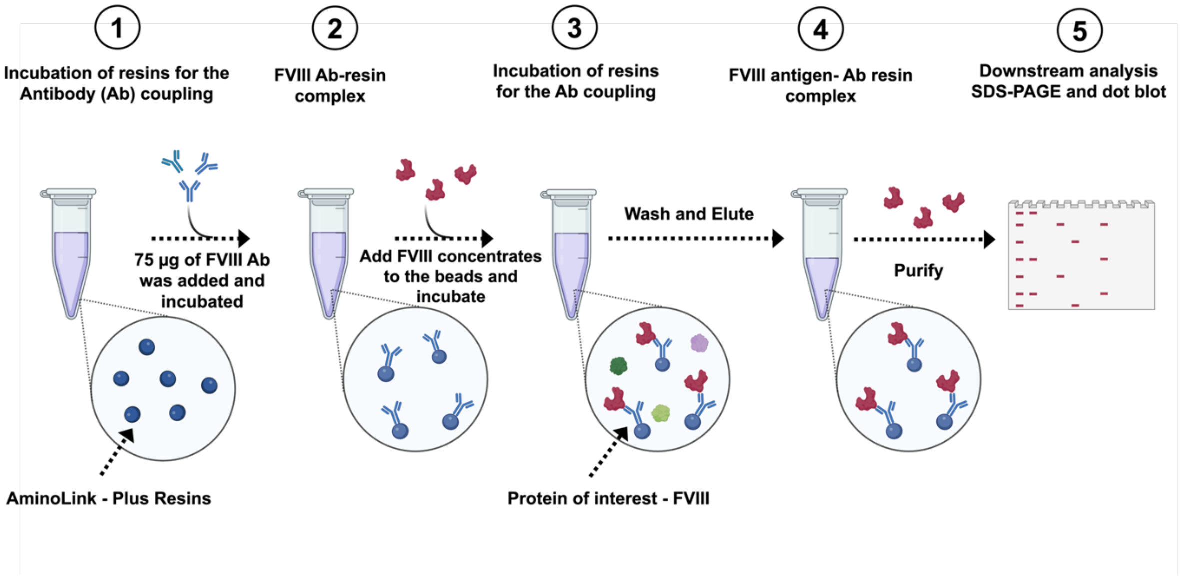


Figure 11. Co-immunoprecipitation purification of FVIII protein. The antibody-based purification of commercially available FVIII products was carried out using a Pierce Co-Immunoprecipitation kit. AminoLink Plus resins were complexed with FVIII and vWF antibodies. The FVIII protein was eluted from the antibody-resin complex and then analyzed for downstream characterization.

- **Ion exchange chromatography (IEX) of rFL-FVIII and pdFL-FVIII:** Purification of rFL-FVIII and pdFL-FVIII proteins via Ion exchange chromatography was performed using Hi-Trap Q-Sepharose Fast Flow column (Q-SP-FF) (Cytiva). Dissociation of the FVIII-vWF complex is usually considered to be tricky as plasma FVIII protein is usually conjugated with vWF protein and its dissociation is influenced by high concentrations of Calcium salt concentrations. [37,130-132] This strategy was also applied to purify rFL-FVIII and pdFL-FVIII proteins. 500 µL of reconstituted proteins was applied onto an anion-exchange resin column. A gradient of calcium chloride ($CaCl_2$ -10-100 mM) was used as the elution buffer to facilitate separation. Before sample application, the column was thoroughly washed with HPLC-grade H_2O and equilibrated with a 1X PBS - calcium-free FVIII buffer. Under these conditions, the neutrally charged FVIII protein bound to the positively charged anion-exchange resin, while impurities and unbound proteins were washed away. The final elution was performed with a high salt concentration of 1M NaCl to remove any unbound FVIII (Figure 12). The washes, FTs, and elutions were then analyzed for highly pure FVIII protein by SDS-PAGE and Dot blot as detailed in Sections 2.3.3.C and 2.3.3.D. Similarly, the pdFL-FVIII protein was purified using mouse anti-human vWF Ab to elute the FVIII proteins from the plasma-derived FVIII-vWF complexes and analyzed by SDS-PAGE with Coomassie staining, and Western blot analysis.

All the proteins purified through SEC, co-IP, and IEX were pooled, and concentrated to obtain high amounts of FVIII protein for further biophysical studies.

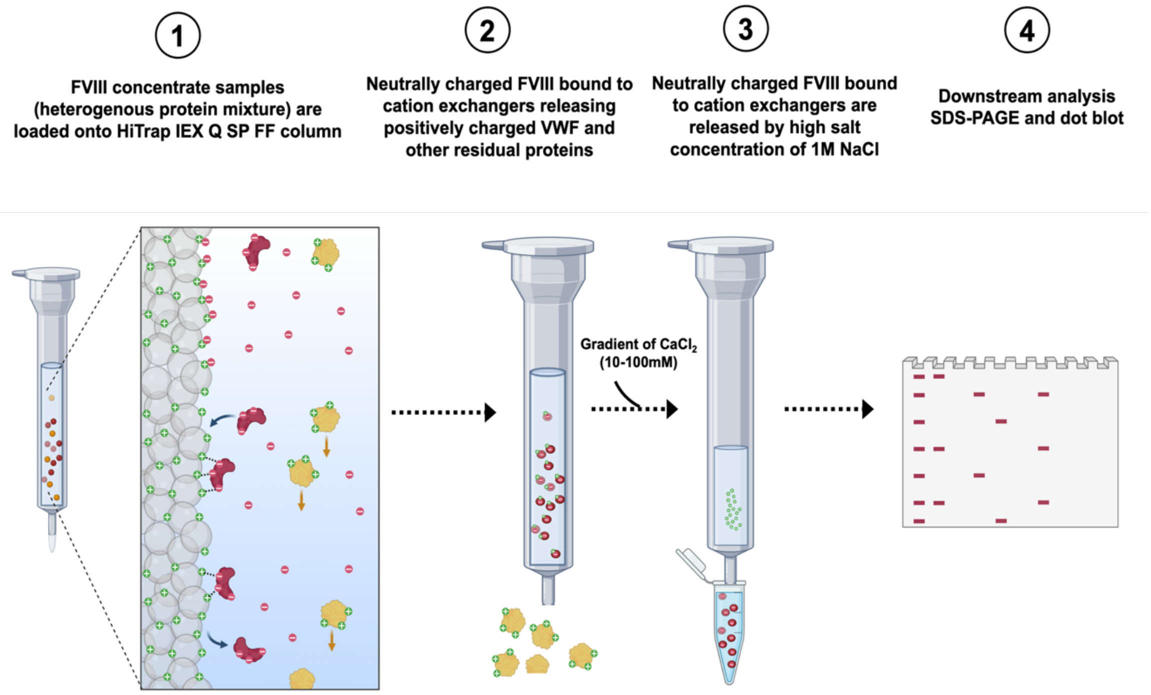


Figure 12. IEX purification of FVIII protein. Anion exchange of FVIII purification was performed using the HiTrap IEX Q SP FF column. The purification was facilitated through selective binding based on charge properties using a gradient concentration of $CaCl_2$-releasing non-targeted proteins. Elution was achieved through a final salt concentration of 1M NaCl, enabling the recovery of the FVIII protein which was further subjected to downstream analysis and further characterization.

2.3.3 FVIII protein biochemistry studies - FVIII protein characterization studies

2.3.3.A FVIII protein characterization studies - Concentration of protein samples

rFL-FVIII, rBDD-FVIII, and pdFL-FVIII purified from SEC, Co-IP, and IEX were concentrated using Amicon® Ultra-0.5 centrifugal 50 kDa molecular weight cut-off (MWCO) filter and Vivaspin® 500 Centrifugal Concentrators - 50 kDa, to obtain fast ultrafiltration and high concentration of highly pure FVIII protein. Considering the FVIII protein size to be ≈ 270 kDa, we utilized 50 kDa MWCO concentrators to eliminate the unwanted residual proteins (albumin) and buffers. The concentrators were spun at 14,000 × g for approximately 10-30 mins at 4 °C until the desired volume/protein concentration was obtained. Highly pure and concentrated FVIII proteins were further utilized for downstream applications.

2.3.3.B FVIII protein characterization studies - Determination of protein sample concentration

- **Nanodrop:** Nanodrop® Spectrophotometer ND-1000 was used to determine the protein concentration by measuring absorbance at 280 nm. To assess potential nucleic acid contamination, absorbance at 260 nm was also measured. 2 µL aliquot of the purified and concentrated rFL-FVIII, rBDD-FVIII, and pdFL-FVIII protein samples was analyzed, with 1X PBS, the buffer in which the proteins were diluted, used as the blank. Protein concentrations of mg/mL were subsequently determined in the Thermo Fischer Scientific Nanodrop 1000 v.3.8.1 software.
- **Bradford assay:** The concentration of purified and concentrated rFL-FVIII, rBDD-FVIII, and pdFL-FVIII protein samples were also determined using the Bradford assay, a colorimetric method based on the binding of Coomassie Brilliant Blue G-250 dye to proteins. A standard curve was prepared using serial dilutions of bovine serum albumin (BSA) stock solution (1 mg/mL) from 0-10 µg/µL. The assay was conducted in a 96-well microplate, where 2 µL of each rFL-FVIII, rBDD-FVIII, and pdFL-FVIII protein sample along with BSA standards were added to the respective wells. Then, 200 µL of Bradford reagent was mixed thoroughly in each well. The plate was incubated at RT for 45 mins to allow the dye-protein complex to form, producing a blue color proportional to the protein concentration. Absorbance was measured at 595 nm using Synergy™ 2 Multi-Mode Microplate Reader embedded with its software BioTek Gen5 microplate reader and Image software v.3.08. Sample protein concentrations were calculated by interpolating their absorbance values against the standard curve

2.3.3.C FVIII protein characterization studies - Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie staining analysis

Protein purity and molecular weight were monitored by analyzing purified rFL-FVIII, rBDD-FVIII, and pdFL-FVIII fractions by SDS-PAGE and subsequent Coomassie-staining. In SDS-PAGE, the proteins were denatured by diluting them in 2X Laemmli sample buffer (Bio-Rad Laboratories, Inc.) containing SDS to a final working concentration of 1X. The SDS, negatively charged masks the charge of the protein itself. Because proteins are denatured and negatively charged, they are separated only according to their relative molecular weight when run on a polyacrylamide gel. Gradient percentage of acrylamide of 4-15 % Mini-PROTEAN® TGX™ Precast Gels were used to run purified and concentrated rFL-FVIII, rBDD-FVIII, and pdFL-FVIII fractions. Approximately 10 µL of the rFL-FVIII, rBDD-FVIII, and pdFL-FVIII fractions (depending on the concentrations calculated as mentioned in the Section 2.3.3.A and Section 2.3.3.B) along with a prestained protein ladder - Precision Plus Protein Dual Color Standards (10 to 250 kDa - Bio-Rad Laboratories, Inc.), electrophoresis was performed at a constant current of 150 V for 45 mins in the Mini-PROTEAN® Tetra cell (Bio-Rad Laboratories, Inc.) using PowerPac™ HC High-Current Power Supply. Coomassie- staining was performed by incubating gels in 0.1 % (weight/volume - w/v) Coomassie R250 1 h at RT on a shaker. Afterward, gels were destained in a destaining solution containing 20 % methanol (volume/volume - v/v) and 10 % glacial acetic acid (v/v) diluted in HPLC-grade H_2O for 2-3 times for 1 h or overnight (o/n). The destained gels were imaged on ChemiDoc™ Touch Imaging System and analyzed further using BioRad Image Lab v.4.1 software. (Bio-Rad). Protein transfer and western blot analysis performed later are described in Sections 2.3.3.C and 2.3.3.D.

2.3.3.D FVIII protein characterization studies - Western and Dot blot analysis

- **Western blot:** Initially, the Western blot analysis was performed using the conventional wet transfer method. 10 µL of FVIII protein samples purified from rFL-FVIII, rBDD-FVIII, and pdFL-FVIII were separated according to their size by SDS-PAGE as described in Section 2.3.3.C. Afterwards the protein was transferred from the gel onto a Polyvinylidene fluoride (PVDF) membrane - Amersham™ Hybond® P Membrane for Western Blot - 0.45 µm. For the transfer, the gel was placed on two layers of filter paper and a sponge previously soaked in 1X Western transfer. The PVDF membrane was activated in 100 % methanol before it was positioned directly on the gel and covered with two more layers of soaked filter paper and a sponge. The transfer was carried out at a constant voltage of 100 V at 4 °C for 90 min. Subsequently, the membranes were blocked for 1 h at RT in 5 % skimmed milk in 1X Tris-buffered saline with 0.1 % Tween (TBS-T). The membranes were incubated with 1:1000 dilution of primary Ab - Sheep anti-Human Factor VIII - PAHFVIII-S FVIII Ab (Haematologic Technologies Prolytix and CellSystems®) and Sheep Anti-Human

Factor VIII, affinity purified IgG - SAF8C-AP FVIII Ab (Affinity™ Biologicals Inc.) 1 h at RT. Primary Ab incubation was followed by 3 washing steps in 1X TBS-T for 10 min each and incubation with 1:4000 secondary Ab Rabbit Anti-sheep IgG antibody, HRP conjugate - Upstate® (Merck-Millipore l Sigma Aldrich Ltd.) - conjugated to horseradish peroxidase (HRP) at RT for 1 h. After incubation with secondary Ab, the membranes were washed 3 times for 30 mins in 1X TBS-T. The membranes were then incubated with equal volumes of Luminol and Peroxidase solution in Immobilon Western HRP substrate (Merck-Millipore GmbH). The images were taken on ChemiDoc™ Touch Imaging System and analyzed further using BioRad Image Lab v.4.1 software. (Bio-Rad). Later the western blot analysis was also carried out by the iBlot™ 2 Gel Transfer device (Invitrogen, Thermo Fisher Scientific™ Inc.) and iBind™ Automated Western System (Invitrogen, Thermo Fisher Scientific™ Inc.)

- **Western blot transfer using iBlot™ 2 Gel Transfer device and Dry Blotting System:** The iBlot™ 2 Gel Transfer device and Dry Blotting System is based on dry blotting, utilizing the unique, patented gel matrix technology developed for the iBlot™ 2 Transfer Stacks, PVDF, mini (Invitrogen, Thermo Fisher Scientific™ Inc.). The iBlot™ 2 Transfer Stack consists of a bottom and a top stack sandwiching a pre-run gel and a PVDF (0.2 µm) membrane. The iBlot™ 2 transfer stacks were assembled with the blotting membrane on the anode side and a pre-run gel on the cathode side. After electrophoresis, the gel was carefully trimmed to fit the transfer stack and positioned directly on the nitrocellulose membrane within the iBlot™ 2 Transfer Stack. Proper alignment using a blotting roller was performed to ensure and avoid air bubbles or wrinkles between the gel and membrane. After the stack was assembled on the iBlot™ 2 Gel Transfer Device and Dry Blotting System, the appropriate method of 20 V for 7 minutes was selected and the run was initiated. Complete transfer of proteins from the gel to the blotting membrane was accomplished in approximately 8 mins.
- **Western blot analysis using iBind™ Western System:** After transferring proteins to a nitrocellulose membrane with the iBlot™ 2 system, the further western blotting procedure was carried out using the iBind™ Western System (Invitrogen, Thermo Fisher Scientific™ Inc.) - an automated western blotting system utilizing sequential lateral flow (SLF) to perform hands-free blocking, antibody binding, and washes for western detection workflows. The iBind™ Western System contains the iBind™ Western device, iBind™ Cards, and iBind™ Solution kit. Initially, the transferred membrane was placed onto the iBind™ Cards, and this card features a specialized matrix designed to facilitate a uniform flow of immunodetection reagents. An iBind™ Solution kit (Invitrogen, Thermo Fisher Scientific™ Inc.), provided with the iBind™ Western System, was used to prepare washing and diluting Ab solutions. To prepare the 1X iBind™ solution for each run, 23.5 mL of iBind™ 5X Buffer was mixed with 500 µL of iBind™ 100X Additive and then diluted to 50 mL with HPLC-grade H_2O.

The iBind™ Card was initially wetted with this solution, and the transferred membrane was placed on top. The lanes of the device were filled in sequence: the first with a primary Ab 1:1000 dilution of primary Ab - Sheep anti-Human Factor VIII - PAHFVIII-S FVIII Ab (Haematologic Technologies Prolytix and CellSystems®) and Sheep Anti-Human Factor VIII, affinity purified IgG - SAF8C-AP FVIII Ab (Affinity™ Biologicals Inc.) in iBind™ buffer, followed by a wash lane with iBind™ buffer to remove excess Ab binding, and then the third lane with a 1:4000 secondary Ab - Rabbit Anti-sheep IgG antibody, HRP conjugate - Upstate® (Merck-Millipore-Sigma Aldrich Ltd.), also diluted in iBind™ buffer, followed by a final wash lane containing the 1X iBind™ solution. The device's automated capillary action and SLF ensured uniform delivery of solutions across the membrane, optimizing Ab binding and washing steps. The membranes were then incubated with equal volumes of Luminol and Peroxidase solution in Immobilon Western HRP substrate (Merck-Millipore GmbH). The images were taken on ChemiDoc™ Touch Imaging System and analyzed further using BioRad Image Lab v.4.1 software. (Bio-Rad).

- **Dot blot:** Additionally, rFL-FVIII, rBDD-FVIII, and pdFL-FVIII proteins purified via co-IP and IEX were mainly analyzed by Dot blot analysis. 2- 5 µL of each sample was directly spotted onto a pre-cut PVDF membrane (Amersham™ Hybond® P Membrane for Western Blot - 0.45 µm). After allowing the spots to air dry completely, the membrane was blocked with 5 % skimmed milk in 1X TBS-T for 1 h at RT to reduce non-specific binding. The membrane was then incubated overnight at 4°C in 1:1000 dilution of primary Ab - Sheep anti-Human Factor VIII - PAHFVIII-S FVIII Ab (Haematologic Technologies Prolytix and CellSystems®) and Sheep Anti-Human Factor VIII, affinity purified IgG - SAF8C-AP FVIII Ab (Affinity™ Biologicals Inc.). Following primary Ab incubation, the membrane was washed three times with 1X TBS-T to remove unbound Abs and then incubated with 1:4000 dilution of secondary Ab - Rabbit Anti-sheep IgG antibody, HRP conjugate - Upstate® (Merck-Millipore l Sigma Aldrich Ltd.) After further washes with 1X TBS-T, the membranes were then incubated with equal volumes of Luminol and Peroxidase solution in Immobilon Western HRP substrate (Merck-Millipore GmbH). The images were taken on ChemiDoc™ Touch Imaging System and dot intensity was analyzed for quantitative comparisons using BioRad Image Lab v.4.1 software. (Bio-Rad). Similarly, dot blot analysis was also performed to check the presence of vWF in the pdFL-FVIII purified samples using 1:1000 dilution of mouse anti-human vWF antibody as the primary antibody. As a secondary Ab, 1:4000 dilution of an anti-mouse IgG antibody was used.

2.3.3.E FVIII protein characterization studies - FVIII mass spectrometry (MS) analysis

To confirm the identity of purified and concentrated rFL-FVIII, rBDD-FVIII, and pdFL-FVIII proteins, samples were analyzed using peptide mass spectrometry. Proteins were enzymatically digested with a protease, such as trypsin, and the resulting peptides were purified and analyzed via mass spectrometry. The peptide masses were then compared to a peptide database to identify the proteins subjected to digestion. For mass spectrometry analysis, proteins were primarily separated by SDS-PAGE as detailed in Section 2.3.3.C. Bands of interest were excised from the gel and transferred into clean 1.5 mL polypropylene tubes. The Coomassie-stained protein bands were destained by incubating the gel pieces with 100 µL of 25 mM ammonium bicarbonate (NH_4HCO_3) and 50 % acetonitrile (ACN) at 37 °C for 15 mins, with repeated destaining until the gel pieces were clear. After removing the destaining solution, 100 µL of 100 % ACN was added for 15 mins at RT to dehydrate the gel pieces. Proteins were then reduced with 10 mM dithiothreitol (DTT) in 100 mM NH_4HCO_3 for 30 mins at 37 °C, followed by alkylation with 55 mM iodoacetamide in 100 mM NH_4HCO_3 for 30 mins at 37 °C in the dark. The gel pieces were washed, dehydrated in 100 % ACN, air-dried, and rehydrated in a trypsin solution (15 ng/µL) for 45 mins on ice. Excess trypsin was removed, the gel pieces were covered with 25 mM NH_4HCO_3, and samples were incubated overnight at 37 °C for digestion. Peptides were extracted twice using 0.1 % formic acid in 60 % ACN, pooled, and acidified with formic acid to a final concentration of 1 % (v/v). The extracted peptides were resolved using a nano-ultra-performance liquid chromatography system coupled to a nano-ESI mass spectrometer (nano Acquity UPLC nanoESI Synapt-MS, Waters, Milford). Separation was achieved using an ACQUITY UPLC M-Class Symmetry C18 Trap Column, 100 Å, 5 µm, and ACQUITY UPLC Peptide CSH C18 analytical column, 130 Å, 1.7 µm, with a 30-minute gradient (3 % to 40 % ACN at 500 nL/min) following three mins of trapping (99 % H_2O at 5 µL/min). Mass spectrometry was performed in V mode, collecting MS^E data using standard parameters. Data analysis employed ProteinLynx Global Server v.2.4, Waters Corporation, and Mascot search engine, referencing an in-house database based on the UniProt database v.2020, restricted to reviewed entries for Homo sapiens; taxon identifier 9606. Protein identifications were accepted with a false positive rate below 4 %. MS analysis was performed at the University of Bonn core facility with the help of Dr. Marc Sylvester and Dr. Robert Hardt.

2.3.3.F FVIII protein characterization studies - FVIII activity assay and vWF multimer analysis

- **Chromogenic Activity Assay for FVIII:** The chromogenic activity of FVIII in the fractions was measured using the Chromogenix Coamatic® Factor VIII kit, commonly used in both clinical and research settings to quantify the procoagulant function of FVIII. Before running the assay, 200 µL of the FVIII purified protein samples were diluted in the buffer working solution provided in the kit to bring the FVIII levels within the dynamic range of the chromogenic assay. The assay reaction involved the use of a chromogenic substrate (S-2765), a synthetic thrombin inhibitor, and a factor reagent containing bovine coagulation factors IXa, X, and thrombin, lyophilized with $CaCl_2$ and phospholipids. In the presence of functional FVIII, FIXa activates FX to Xa, which subsequently hydrolyses the chromogenic substrate, releasing para-nitroaniline (pNA), a yellow-colored compound. After the reaction components were incubated, the absorbance of the yellow color was measured spectrophotometrically at 405 nm. The intensity of the color is directly proportional to the amount of functional FVIII in the sample, as more FVIII leads to more FXa, which in turn increases the hydrolysis of the substrate. Absorbance readings were calibrated using a standard curve generated from known concentrations of FVIII provided by the kit's reference standards. Data were plotted to evaluate FVIII activity at each time point, allowing for the comparison of FVIII variants or different experimental conditions. Results were expressed in international units (IU) of FVIII activity per mL, enabling a clear comparison of activity levels over time.
- **vWF Multimer Analysis:** Analysis of vWF multimers was done using the Hydragel-5 von-Willebrand multimers kit (vWF-MM kit) run on the semiautomated Hydrasys-2 Scan Focusing system according to the manufacturer's instructions (Sebia, Mainz, Germany). In brief, total vWF-antigen (Ag) in the single fractions of pdFL-FVIII was determined by applying the immunoturbidimetric vWF-Ag assay run on a CN-6000 coagulation analyzer (Siemens Healthineers, Erlangen, Germany). Subsequently, samples were diluted according to the specifications of the vWF-MM kit and analyzed in parallel to an in-house normal pooled citrated plasma (NPP) (prepared from the donations of 4 healthy blood donors). Analysis of the relative distribution of vWF multimers in the purified peak fractions of pdFL-FVIII was done using Phoresy's Core software (Sebia).

2.3.3.G FVIII protein characterization studies - CD studies

Far-ultraviolet (UV) circular dichroism (CD) measurement of protein secondary structure was conducted using a Jasco-J715 spectrophotometer (JASCO, USA). Wavelength scans in the far-UV regions of 180-300 nm were performed in 0.1 cm path-length Quartz glass cuvette at RT. The scanning was performed in a continuous mode with a scanning rate of 100 nm/min and a bandwidth of 1 nm. 200 µL of purified rFL-

FVIII and rBDD-FVIII samples were utilized for a direct quantification of protein amount using Nanodrop® Spectrophotometer ND-1000 prior to CD measurements. Baseline correction with buffer was performed before loading the proteins. For each sample, five spectral measurements were performed. The data were analyzed on the BeStSel webserver and DichroWeb. [133-136]

2.3.3.H FVIII protein characterization studies - Deglycosylation studies

Deglycosylation studies were performed initially with an aim to perform cross-linking mass spectrometry studies (XL-MS), while additionally, it was also performed to understand the role of glycosylation in maintaining the stable FVIII protein. Deglycosylation of purified rFL-FVIII proteins was performed through PNGase F digestion according to the manufacturer's instructions (New England Biolabs - NEB). Approximately, 5 μg of rFL-FVIII purified protein was dissolved in a 1X Glyco-denaturing buffer and heated up at 100°C for 10 mins for denaturation of the protein. Later, the denatured protein was mixed with GlycoBuffer 2, and PNGase F and incubated at 37°C for 1 hour. The PNGase F cleaves between the innermost GlcNAc and Asn residues of high mannose, hybrid, and short complex oligosaccharides. Removal of the glycans led to a change in the MW of the rFL-FVIII protein causing a mobility shift which was analyzed by SDS PAGE and coomassie staining and further analyzed using mass spectrometry for the deglycosylated residues as mentioned in Sections 2.3.3.C and 2.3.3.E.

2.4 Methods - Biophysical studies for structural characterization of FL-FVIII and FL-FVIII-vWF complex.

2.4.1 Biophysical studies - Negative staining and Cryo-Electron Microscopy studies

- **Negative staining studies:** The stains used for the process are the Sodium Silico Tungstate (SST - $Na_4O_{40}SiW_{12}$) at 2 % in distilled water (pH 7 - 7.5) Uranyl Acetate (AcUr) $UO_2(CH_3COO)_2.2H_2O$ at 2 % in distilled water (pH 4.2 - 4.5). A negative Stain-Mica-carbon Flotation Technique (MFT) was performed to image the purified peak-1 rFL-FVIII and peak-2 of pdFL-FVIII proteins. A 5 μL sample was applied to the clean side of a carbon film on mica, stained, and transferred onto a 400-mesh copper grid. Images were captured under low-dose conditions (<10 $e^-/Å^2$) with defocus values ranging from 1.2 to 2.5 μm using a Tecnai 12 LaB6 electron microscope (EM) and at an accelerating voltage of 120 kV with a 4K Field Electron and Ion (FEI) Eagle Charge-Coupled Device (CCD) camera. Image processing was carried out in RELION v.3.1.2, and contrast transfer function (CTF) estimation was performed with CTFFind-4.1. An initial particle set (box size 256 pixels, sampling 2.2 Å/pixel) was obtained via autopicking. A second set of particles (box size 256 pixels, sampling 4.4 Å/pixel, mask diameter 180 Å) was also collected using a

Gaussian blob auto-picking method. Following 2D classification, the most distinguishable 2D class averages were selected for further analysis, and final classifications were performed. The final classification was calculated. Initial rFL-FVIII negative staining studies were performed as a part of a collaboration with Prof. Dr. Guy Schoen and Dr. Jean-Luc Pellequer at the University of Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale (IBS), F-38000 Grenoble, France. Additionally, negative staining studies of rFL-FVIII and pdFL-FVIII proteins were also repeated as a part of collaboration with Dr. Monika Gunkel and Prof. Elmar Behrmann at the StruBiTEM (Structural Biology Cryo-Transmission Electron Microscopy), Institute of Biochemistry, University of Köln. A similar procedure was performed to image the purified peak-1 rFL-FVIII and peak-2 of pdFL-FVIII proteins. A 5 µL sample was applied to the clean side of a carbon film on mica, stained, and transferred onto a 400-mesh copper grid. Images were captured under low-dose conditions (<24 e^-/Å^2) with defocus values ranging from -0.5 to -2.5 µm using a Talos L120C electron microscope (EM) and at an accelerating voltage of 120 kV with a Ceta camera. Image processing was carried out in RELION v.3.1.2, and CTF estimation was performed with CTFFind-4.1. An initial particle set (box size 256 pixels, sampling 1.89 Å/pixel) was obtained via auto picking. A second set of particles (box size 256 pixels, sampling 4.4 Å/pixel, mask diameter 180 Å) was also collected using a Gaussian blob auto-picking method. Following 2D classification, the most distinguishable 2D class averages were selected for further analysis, and final classifications were performed (Figure 13, Table 2).

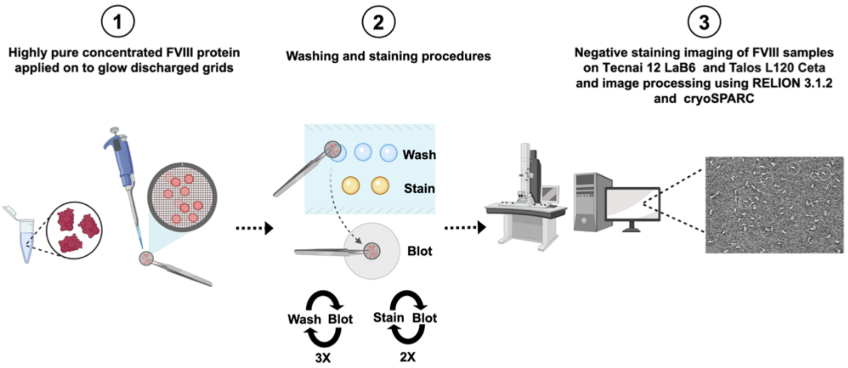


Figure 13. Negative staining of rFL-FVIII and pdFL-FVIII proteins. Highly purified, concentrated FVIII protein is applied onto glow-discharged grids to ensure uniform protein adsorption. The grids were allowed for sequential washing and staining steps to remove unbound protein and achieve optimal contrast for imaging. The negatively stained FVIII samples are imaged using Tecnai T12 LaB6 and Talos L120C Ceta microscopes. Image data are processed using RELION v.3.1.2 and cryoSPARC software to obtain negative topographs of the FVIII protein.

	rFL-FVIII	rFL-FVIII	pdFL-FVIII
Microscope	Tecnai 12 LaB6 electron	Talos L120C	Talos L120C
Camera	4K FEI Eagle CCD	Ceta	Ceta
Accelerating voltage [kV]	120	120	120
Spherical Aberration Cs [mm]	2.7	2.7	2.7
Size of objective aperture [µm]	No aperature	No aperature	No aperature
Size of Condenser 2 (C2) aperture [µm]	50	50	50
pixel size [Å/px]	Other	Other	Other
Magnification	Other	Other	Other
exposure time [s]	0.5	0.5	0.5
Total exposure dose [$e^-/Å^2$]	10	24	24
Images per hole	1	1	1
defocus range start value in µm	1.2	-0.5	-0.5
defocus range end value in µm	1.5	-2.5	-2.5
number of images acquired	100	100	100
Other Magnification	73000	73000	73000
Other pixel size [Å /px]	2.2	1.89	1.89

Table 2. Negative staining parameters and details of rFL-FVIII and pdFL-FVIII proteins.

- **Cryo-electron microscopy studies :**

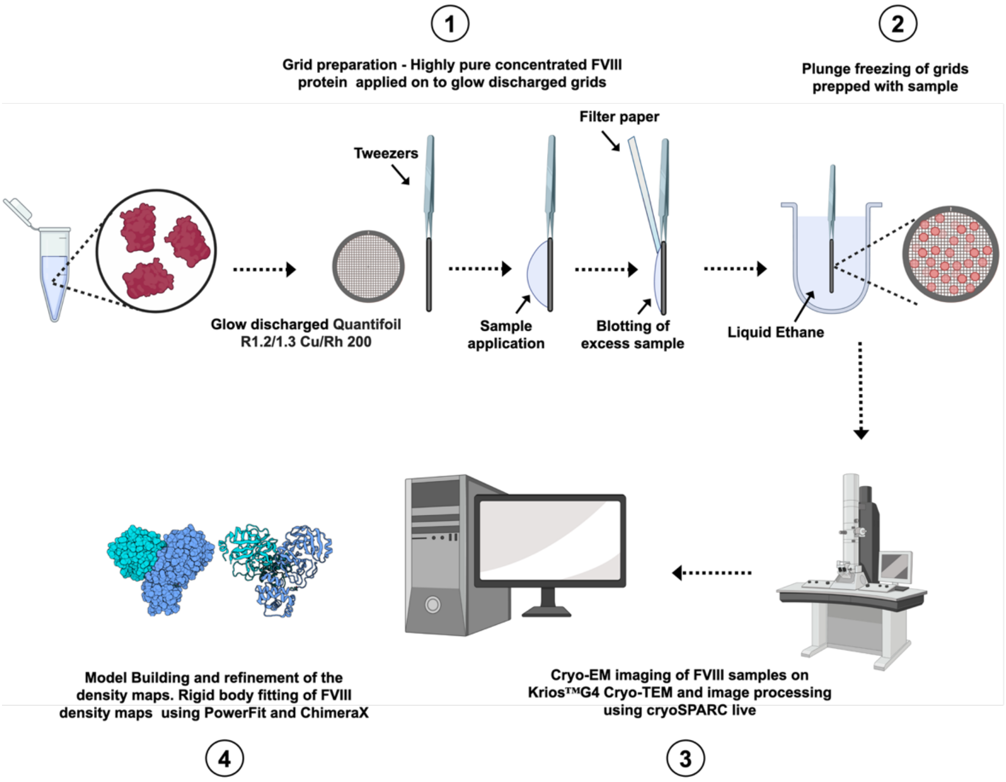


Figure 14. cryo-EM analysis of rFL-FVIII and pdFL-FVIII proteins. The cryo-EM analysis of FVIII protein was initiated by applying highly pure, FVIII proteins onto glow-discharged Quantifoil R1.2/1.3 Cu/R200 grids. The excess sample was blotted using filter paper to achieve an even distribution. Plunge freezing of the prepared grids is rapidly frozen in liquid ethane and Vitrobot mark IV to vitrify the sample. Data collection was performed on Krios™ or Talos™ Cryo-TEM systems, and image processing was carried out using cryoSPARC live for high-resolution structural reconstruction. Density maps obtained were refined, and rigid body fitting of FVIII structures was performed using tools like PowerFit and ChimeraX to elucidate the structural and functional details of the FVIII protein.

- **Recombinant FL-FVIII:** Approximately 3µL of purified rFL-FVIII was loaded onto EM-grids (Quantifoil® R1.2/1.3 200 Mesh, Cu/Rh) that were glow discharged (20 mAmp, 45 s, Quorum GloQube, Quorum Technologies), incubated for 45 s and blotted for 5-7 s, force 0, at 100 % humidity and 4 °C, before plunge freezing into liquid ethane using a FEI Vitrobot mark IV plunge freezer (Thermo Fisher Scientific™ Inc.). For dataset acquisition, the grids were screened and analyzed using Thermo Fisher Scientific™ Inc. Glacios Cryo-Transmission electron microscopy (cryo-TEM). Then the grids were brought to Thermo Fisher Scientific™ Inc. Krios™ G4 Cryo-TEM equipped with an E-CFEG, a Thermo Fisher Scientific™ Inc. Selectris™ X Energy Filter, and a Falcon 4 Detector. Movies were collected with Aberration-free image shift (AFIS), with a pixel size of 0.93 Å, a defocus range of -0.9 to -1.8 µm, and a total dose of 40 $e^-/Å^2$. Cryo-EM data were processed with cryoSPARC v.3.3.2. The movies were first motion-corrected and a patch-based CTF estimation was then performed using cryoSPARC live. After patch-motion and CTF correction, particles were picked which were then subjected to 3 rounds of 2D classification. The remaining particles were used for Ab-initio reconstruction and 3D refinement.

- **Plasma-derived FL-FVIII:** 3 µL purified pdFL-FVIII sample is applied to the grids (Quantifoil® R1.2/1.3 200 Mesh, Cu) that were glow discharged (20 mAmp, 40 s, Quorum GloQube, Quorum Technologies). The grids were plunge frozen in liquid ethane using an FEI Vitrobot Mark IV plunge freezer (Thermo Fisher Scientific™ Inc.) at 4°C and >90 % humidity with 5-7 s blot time and 0 blot force. The grids were screened and analyzed using a Thermo Fisher Scientific™ Inc. Krios™G4 Cryo-TEM equipped with an E-CFEG, a Thermo Fisher Scientific™ Inc. Selectris™ X Energy Filter, and a Falcon 4 Detector. Movies were collected with a pixel size of 0.727 Å and a total dose of 50 $e^-/Å^2$. Cryo-EM data were processed with cryoSPARC v.3.3.2. The movies were first motion-corrected and a patch-based CTF estimation was then performed using cryoSPARC live. Particles were picked using the blob-picking algorithm. 3 rounds of 2D classification and further particle analysis were used for Ab-initio reconstruction with four classes. The first particle class was subjected to and 3D Homogenous Refinement for final 3D reconstruction (Figure 14, Table 3). The cryo-EM studies mentioned here were performed as a part of a collaboration with Ms. Boxue Ma, and Ms. Deniz Ugurlar, from ThermoFisher Scientific, Cryo-EM Unit, Eindhoven, Netherlands.

	rFL-FVIII and pdFL-FVIII
Concentration [mg/ml]	1
Sample volume [µl]	3
Grid type	Quantifoil R1.2/1.3 Cu/Rh 200 and Quantifoil R1.2/1.3, 200 respectively
Glow discharge time [s]	45
Glow discharge current [mA]	20
Blotting chamber temperature [°C]	4
Blotting chamber humidity [%]	100
Blot time [s]	5
Voltage	300V
TEM mode	EER
Indicated magnification	130

Target defocus [µm]	0.6 - 2
Zero-loss slit width [eV]	10
Pixel size [Å]	0.93
Total exposure [e/Å^2]	40

Table 3. cryo-EM parameters and details of rFL-FVIII and pdFL-FVIII proteins.

2.4.2 Biophysical studies - Atomic Force Microscopy studies

- **Air AFM:** AFM imaging of rFL-FVIII and rBDD-FVIII samples in the air was performed using high-resolution probes to capture the topographic details. Two types of probes were employed, depending on the imaging requirements. Triangular Scan-Asyst probes with a spring constant (k) of 0.4 N/m, a resonance frequency (F_0) of 70 kHz, a length (L) of 115 µm, and a tip radius of 2 nm were used to capture broader surface details, while rectangular Scan-Asyst-HR probes with the same spring constant, a higher resonance frequency of 130 kHz, and a shorter length of 60 µm were used for more precise, high-resolution imaging. The AFM was operated in air, allowing for higher resolution imaging by minimizing thermal noise. 2 µL of each rFL-FVIII and rBDD-FVIII samples were deposited on freshly cleaved muscovite mica (BrukerAFMProbes, Camarillo, Santa Barbara). After the sample was deposited, it was left to dry in air at room temperature for 10-15 mins. To speed up the drying process, sometimes the samples were also dried under a gentle stream of nitrogen. The drying process allowed any residual solvent or buffer to evaporate without causing aggregation or denaturation of the proteins. Using ScanAsyst mode, the imaging parameters, including force control and scan speed, were optimized automatically, reducing potential damage to the protein structures. The sharp probe tips, with a 2 nm radius, allowed for the resolution of nanoscale topographic features of these samples, yielding FL-FVIII and BDD-FVIII air AFM topographs (Figure 15).
- **Liquid AFM:** AFM imaging of rFL-FVIII and rBDD-FVIII samples in liquid was performed on freshly cleaved muscovite mica (BrukerAFMProbes, Camarillo, Santa Barbara). 1 µL of each protein sample was deposited on the mica and immediately covered with 50 µL of pure water to maintain hydration and preserve the protein's native structure during imaging. AFM images were acquired in a liquid environment using Triangular ScanAsyst-Fluid probes (k = 0.7 N/m; F0 = 150 kHz; L = 70 µm; tip radius 20 nm; BrukerAFMProbes, Camarillo), making them suitable for accurately capturing surface features without excessive sample deformation. The PeakForce tapping mode was employed in a semi-automated setup, allowing for controlled interaction between the AFM tip and the sample surface while adjusting parameters such as ramp size (150 nm) and manual gain adjustments to optimize image quality. Additionally, The AFM imaging and further image processing of the Laplacian weight

filter of AFM images were obtained on a multimode 8-nanoscope V controller (BrukerAFMProbes, Camarillo, Santa Barbara) which enabled detailed topographical analysis of the protein samples in their hydrated state. Similarly purified peak-2 purified pdFL-FVIII-vWF complex, the DriveAFM system (Nanosurf AG, Switzerland) was utilized. This AFM system employs a specialized cantilever, the WM0.6 AuD (k = 0.6 N/m, F0 = 350 kHz, L = 35 µm, with a tip radius of 10 nm), designed for photothermal excitation. This yielded FL-FVIII, BDD-FVIII, and FL-FVIII-vWF liquid AFM topographs (Figure 15).

- **Image Processing:** FL-FVIII, BDD-FVIII, and FL-FVIII-vWF AFM topographs were processed and corrected using Gwyddion. [137] Initially, the original AFM height of all the images was flattened using Gwyddion while stripe noise was removed with DeStripe when needed. [138] The loaded topographs were primarily examined for artifacts such as tilt, noise, or scan line distortions. Briefly, a plane fit correction was applied to the background to level the sample, ensuring the average height across the image was neutralized, followed by a line correction using the median option. Scan line artifacts were corrected using the Align Rows function, which adjusted misaligned rows. Surface flattening was performed using Polynomial Flattening. Noise reduction was achieved by applying Gaussian filters, with filter radius adjustments to preserve topographic features while eliminating high-frequency noise. Strip noise was removed using DeStripe. For region-specific analysis, the area of interest for individual topographs was cropped using the selection tool and focused on relevant regions. [86,87,89] Cropped topographs were then selected using the thresholding method and a new line correction was applied using a third-order polynomial fit with selected grains excluded (Figure 15). Finally, the corrected and cropped topographs were utilized for downstream DockAFM studies mentioned in Section 2.5.2. The entire AFM studies were carried out as part of a collaborative initiative during Ms. Urs's visiting research experience, supported by the Rudolf Marx Stipendium, in collaboration with Dr. Jean-Luc Pellequer at the University of Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale (IBS), F-38000 Grenoble, France. Also, the AFM images of the FL-FVIII-vWF complex in a liquid environment were obtained during the 14th AFM BioMed Summer School at Marcoule, where access to the DriveAFM system was provided for FVIII-vWF complex analysis. This work was carried out with the invaluable assistance of Dr. Patrick Frederix from Nanosurf AG, Liestal, Switzerland.

2.5 Integrating biophysical studies data and computational models

2.5.1 Integrating biophysical studies data and computational models - Fitting computational models to low-resolution cryo-EM density maps

Rigid body fitting of the refined *Gly/FC-FL-FVIII* model to both rFL-FVIII and pdFL-FVIII cryo-EM density maps was performed using the PowerFit server (https://alcazar.science.uu.nl/cgi/services/POWERFIT/powerfit/submit), a computational tool designed to identify the optimal placement of atomic structures within cryo-EM density maps and Fit in Map function of ChimeraX. [85,139-142] The refined *Gly/FC-FL-FVIII* model PDB file and both rFL-FVIII and pdFL-FVIII cryo-EM density maps in the Medical Research Council (MRC) format were uploaded to the PowerFit server. The cryo-EM maps were pre-processed by applying appropriate filters, such as low-pass filtering in ChimeraX, to match the resolution of the atomic model and reduce noise. Once uploaded, the server parameters were configured. These parameters included the map resolution, which was specified based on the cryo-EM map's Fourier Shell Correlation (FSC) cut-off, and the grid sampling rate for computational efficiency. PowerFit uses a multi-scale approach to scan the entire density map with the atomic model and calculates the cross-correlation score (CCS) between the atomic model and the density map to determine the best-fitting of the HC $_{\text{(A1-A2-domains)}}$, B-domain, and LC $_{\text{(A3-C1-C2-domains)}}$ of *Gly/FC-FL-FVIII* model at various orientations and positions. After the computation, the output provided the top-ranked poses based on their CCS, indicating the optimal fit of the atomic structure within the map. The best fit was visually inspected and further refined using molecular visualization software, such as ChimeraX, to confirm alignment and ensure compatibility with the biological context. The Fit_in_Map in ChimeraX command was then executed, specifying the model and map identifiers. This function optimized the positioning by adjusting the translation and rotation of the HC $_{\text{(A1-A2-domains)}}$, B-domain, and LC $_{\text{(A3-C1-C2-domains)}}$ of the *Gly/FC-FL-FVIII* model to maximize its correlation with the density observed in rFL-FVIII and pdFL-VIII cryo-EM maps. The refined placement of the model within the map and a computed correlation score indicated the fit quality.

2.5.2 Integrating biophysical studies data and computational models - Fitting computational models to AFM topographs through AFM Assembly and DockAFM pipelines

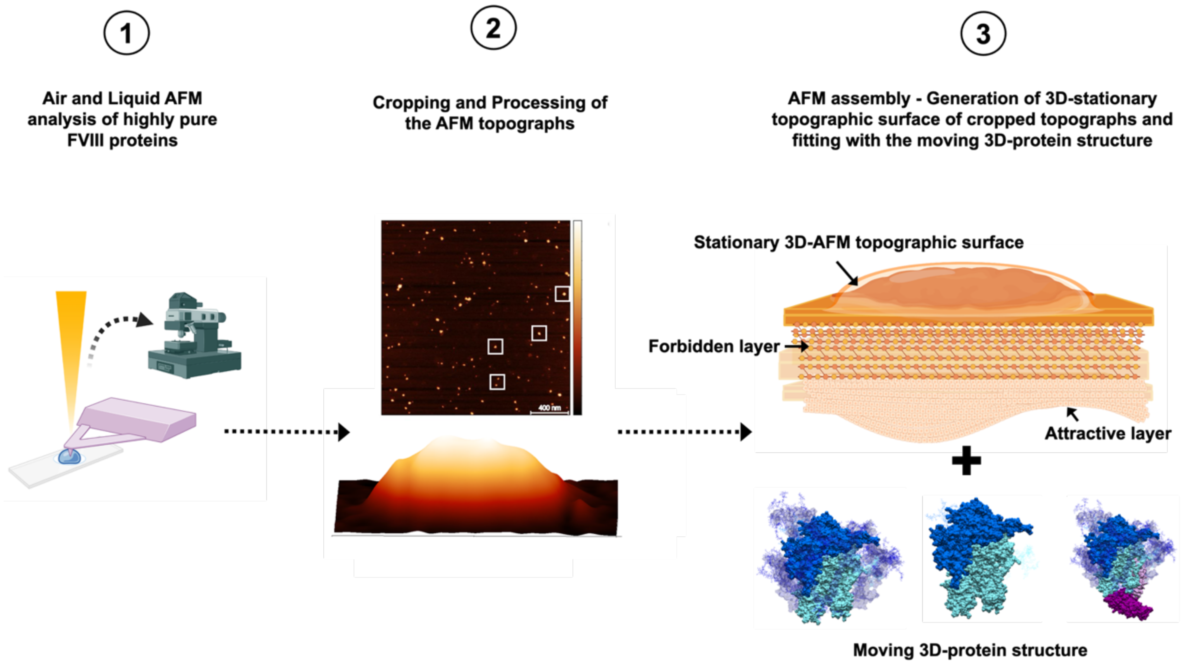


Figure 15. AFM and DockAFM analysis of the rFL-FVIII, rBDD-FVIII, and pdFL-FVIII proteins. Both air and liquid AFM analyses of highly pure rFL-FVIII, rBDD-FVIII, and pdFL-FVIII proteins were analyzed. Further, image analysis, cropping, and processing were performed using Gwyddion and the cropped topographs were subjected to AFM assembly using DockAFM. The stationary 3D-AFM topographic surface was generated and combined with the moving 3D-protein structure through fitting, incorporating the attractive and forbidden layers to provide a comprehensive model of the FVIII protein's topography and dynamics.

The original height of all the FL-FVIII, FL-FVIII-vWF, and BDD-FVIII AFM topographs were flattened and denoised using Gwyddion mentioned in Section 2.4.2. The de-noised images were cropped and adjusted using Gwyddion, and an exhaustive search for all possible docking orientations was carried out using the DockAFM/AFM Assembly pipeline. [88,143] This pipeline was updated to the 2024 version using a newer version of DOT2, an FFT-based rigid docking software. DOT2 software was used to translate denoised and cropped FL-FVIII, FL-FVIII-vWF, and BDD-FVIII images into volumetric data. Further, this was rigorously utilized for a six-dimensional search (translation in X, Y, and Z axes and three Euler rotations) to fit our respective models. The DOT2 algorithm was also used to translate and rotate 3D atomic coordinates about a stationary object. Using the DOT2 software, each protein fragment undergoes exhaustive rigid-body fitting through translation and rotation across the grid. Here, the stationary objects were the AFM topographs (height image) and the moving object was the surface layer corresponding to each molecular model and their predicted dockings (Figure 15). For this study, comparative DockAFM studies were performed as mentioned below.

- **FL-FVIII assembly and docking:** Cropped FL-FVIII AFM topographs were fitted and docked with refined and validated *Gly/FC-FL-FVIII* model *and AF-FL-FVIII* model. Here, the FL-FVIII cropped topograph was considered as a stationary object which was adjusted to a resolution of 0.5-1 nm/pixel and their respective computational refined models of *Gly/FC-FL-FVIII and AF-FL-FVIII* were moving objects to perform optimal docking/fitting. The FL-FVIII AFM topographic surface was interpolated to match grid pixel counts, and data points were classified as "forbidden" or "attractive" based on their position relative to the surface. Specific parameters for FL-FVIII AFM topographic surface included a 128 cubic grid with a grid spacing of 4.9 Å in XY and 4 Å in Z; a baseline threshold of 30 Å; a forbidden layer of 25 Å and a favorable layer of 40 Å; a rotation step of 6°. The models were then subjected docking process yielding the results ordered according to a pseudo-energy score at a given XYZ offset in the grid. The comparative studies were performed to obtain the best docking orientation with the lowest energy as well as a quality check of the models and AFM topographs.
- **BDD-FVIII assembly and docking:** Cropped BDD-FVIII AFM topographs were fitted and docked with refined and validated *Gly-BDD-FVIII* and *Gly/FC-FL-FVIII* models. Here, the BDD-FVIII cropped topograph was considered as a stationary object which was adjusted to a resolution of 0.5-1 nm/pixel and their respective computational refined models of *Gly-BDD-FVIII and Gly/FC-FL-FVIII* were moving objects to perform optimal docking/fitting. The BDD-FVIII AFM topographic surface was interpolated to match grid pixel counts, and data points were classified as "forbidden" or "attractive" based on their position relative to the surface. Specific parameters for BDD-FVIII AFM topographic surface included a 128 cubic grid with a grid spacing of 4.9 Å in XY and 4 Å in Z; a baseline threshold of 30 Å; a forbidden layer of 25 Å and a favorable layer of 40 Å; a rotation step of 6°. The models were then subjected docking process yielding the results ordered according to a pseudo-energy score at a given XYZ offset in the grid. The comparative studies were performed to obtain the best docking orientation with the lowest energy as well as a quality check of the models and AFM topographs.
- **FL-FVIII-vWF assembly and docking:** Cropped FL-FVIII-vWF AFM topographs were fitted and docked with refined and validated *Gly/FC-FL-FVIII-D'-D3 complex* model and *Gly/FC-FL-FVIII* model. Here, the FL-FVIII-vWF cropped topograph was considered as a stationary object which was adjusted to a resolution of 0.5-1 nm/pixel and their respective computational refined models of *Gly/FC-FL-FVIII-D'-D3 complex and Gly/FC-FL-FVIII* were moving objects to perform optimal docking/fitting. The FL-FVIII-vWF AFM topographic surface was interpolated to match grid pixel counts, and data points were classified as "forbidden" or "attractive" based on their position relative to the surface. Specific parameters for FL-FVIII-vWF AFM topographic surface included a 128 cubic grid with a grid spacing of 4.9 Å in XY and 4 Å in Z; a baseline threshold of 30 Å; a forbidden layer of 25 Å and a

favorable layer of 40 Å; a rotation step of 6°. The models were then subjected docking process yielding the results ordered according to a pseudo-energy score at a given XYZ offset in the grid. The comparative studies were performed to obtain the best docking orientation with the lowest energy as well as a quality check of the models and AFM topographs.

The DockAFM studies were carried out as part of a collaborative initiative during Ms. Urs's visiting research experience, supported by the Rudolf Marx Stipendium, in collaboration with Dr. Jean-Luc Pellequer at the University of Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale (IBS), F-38000 Grenoble, France.

2.6 Methods - Functional characterization and interaction studies

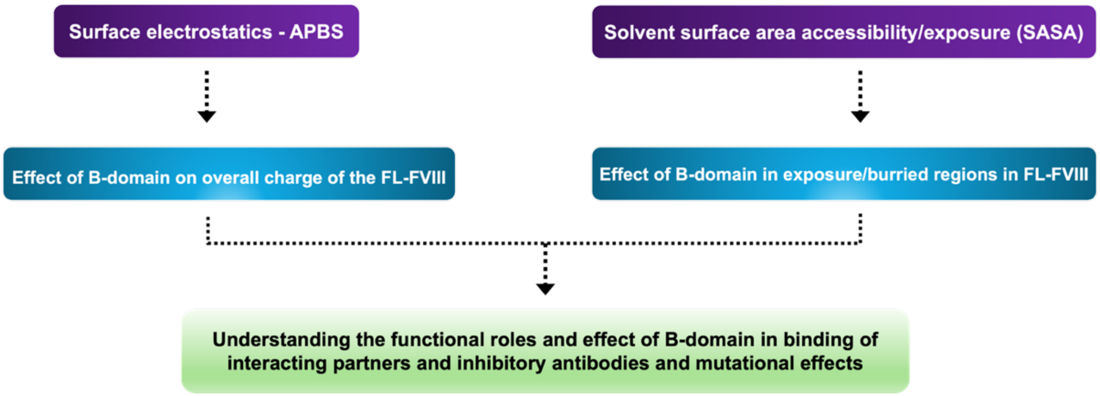


Figure 16. Schematic representation of the functional characterization and interaction studies of B-domain. The left panel highlights the use of surface electrostatics analysis (APBS) to evaluate the impact of the B-domain on the overall charge of FL-FVIII. The right panel demonstrates the assessment of solvent surface area accessibility/exposure (SASA) to study the role of the B-domain in exposed and buried regions of FL-FVIII. Together, these analyses were used to understand the functional roles of the B-domain in binding interacting partners, inhibitory antibodies, and its mutational effects.

2.6.1 Functional characterization studies - Electrostatic studies

Electrostatic analysis of the *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII* models was conducted using PyMOL integrated with Adaptive Poisson-Boltzmann Solver (APBS) plugin. [144,145] Before calculations, the models were prepared using PDB2PQR to assign atomic charges, optimize hydrogen bonding, and adjust titratable groups at a physiological pH of 7.4, ensuring an accurate representation of their chemical properties. The prepared files were then used for Poisson-Boltzmann calculations, with parameters set to simulate physiological conditions, including a solute dielectric constant of 2, solvent dielectric constant of 80, and ionic strength of 0.15 M. The calculated electrostatic potentials were mapped onto the solvent-accessible surface of the models and visualized in PyMOL, with a color-coded scheme representing positive, negative, and neutral potentials. This analysis was used to reveal individual regions of electrostatic significance, such as charged clusters in *Gly/FC-FL-FVIII* HC $_{\text{(A1-A2-domains)}}$, LC $_{\text{(A3-C1-C2-domains)}}$, and B-domain. These charged clusters were compared to the ones in *Gly-BDD-FVIII* HC $_{\text{(A1-A2-domains)}}$, and LC $_{\text{(A3-C1-C2-domains)}}$. The resulting data were exported for further analysis to understand the electrostatic landscape and columbic potential differences between *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII*. The differences were also used to analyze FL-FVIII's structure and the B-domain's functional roles in binding to PS membrane residues - R2182 and R2339 (Figure 16).

2.6.2 Functional characterization studies - Surface area accessibility studies

The solvent-accessible surface area (SASA) analysis provided a quantitative measure of the surface area of *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII* models that is accessible/buried to a solvent molecule in the presence and absence of B-domain and glycosylations. SASA of the refined models was determined using YASARA and MSMS with a threshold of 1 Å. [146] The analysis involved the rolling probe method, where a spherical probe of 1.4 Å radius, simulating the size of a water molecule, was rolled over the protein surface to calculate accessible regions. Before the calculation, both the *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII* models were optimized within YASARA to minimize steric clashes and ensure accurate atomic positioning. The SASA tool in YASARA was then employed to calculate residue-specific and total solvent-accessible surface areas. The analysis also incorporated atomic radii and standard solvent parameters, offering a detailed perspective on the distribution of hydrophilic and hydrophobic regions across the models. The results included residue-level SASA values in A^2 and later were mapped onto the models for visualization, highlighting exposed and buried regions. Further, the total SASA values of the *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII* models were compared by calculating the difference between them. Since the primary distinction between these two models lies in the presence of the B-domain and additional glycosylations, the observed differences were attributed to these features (Figure 16). This SASA variation was subsequently leveraged to conduct functional studies, aiming to investigate the impact of the B-domain and glycosylations on the binding of inhibitory antibodies, interactions with protein partners, and the role of mutations associated with HA (as detailed below)

2.6.2.A Surface area accessibility studies: FVIII inhibitory antibodies

- **Inhibitors:** Our functional characterization studies mainly included calculating the SASA of inhibitory antibody epitopes found to be bound to the A2-, A3-, C1-, and C2-domains of FVIII. Both classical and non-classical neutralizing inhibitory antibodies (NAbs) with/without available crystal structures were utilized to understand the effect of B-domain and glycosylations on the major surface epitopes involved. The crystal structures including classical and non-classical inhibitor complexes were retrieved from the RCSB-PDB and information regarding all the epitopes was also extracted from relevant literature sources. SASA of these major epitopes involved (mentioned below) were calculated between the refined *Gly/FC-FL-FVIII and Gly-BDD-FVIII* models and compared. Additionally, tools such as the SpotOn webserver and PDBePISA (Proteins, Interfaces, Structures, and Assemblies) were employed to identify further interface contacts and assess interactions at the atomic level.

  Classical inhibitory antibodies included - **PDB-ID: 4XZU** binary C2/3E6 complex with epitopes in C2-domain - G2198-D2207, T2210-S2212, S2223-R2234, Q2241-W2248, D2252, L2280-S2282, G2287-V2301, R2326-Q2330. **PDB-ID: 7K66**, an ET3i/2A9 complex with epitopes in C2-domain - G2198-D2207, T2210-S2212, S2223-R2234, Q2241-W2248, D2252, L2280-S2282, G2287-V2301, R2326-Q2330. **PDB-ID: 1IQD (NB2C11)**, a C2/ G4$_k$ Fab complex (BO2C11 F$_{AB}$), with epitopes in C2-domain - F2215-A2220, G2233-R2239, Q2241-N2243, S2269-M2274, H2334-Q2335. **PDB-ID: 8G6I (NB33)**, a recombinant derivative of KM33 with epitopes in C1-domain - R2109-S2113, I2177-R2178.

  Non-classical inhibitory antibodies included - **PDB-ID: 4KI5**, C2/3E6/G99 F$_{AB}$ ternary complex with epitopes in C2-domain - G2198-A2207, S2212-S2213, T2221-R2234, Q2241-W2247, L2280-S2282, W2290-L2292, K2298-V2301, R2326-Q2330, R2339. **PDB-ID**: **8TY1 (NB2E9)**, with C1-domain epitopes - S2059-Y2062, K2084-W2089, R2169-H2174. **NBG12**, with C1-domain epitopes - A2096-I2103. **NBA3**, with A3-domain epitopes - K1713-Y2036. **NB41**, with A3-domain epitopes - R1822-K1837, and **NB11B2**, with A2-domain epitopes - P398-E475.

The analysis of these surface epitopes involved detailed mapping of the solvent accessibility of key residues within the inhibitor-bound interfaces. The comparison of SASA values between *Gly/FC-FL-FVIII and Gly-BDD-FVIII* models helped to elucidate how these antibodies exert inhibitory effects by targeting specific epitopes within the A2-, A3-, C1-, and C2-domains. This expanded dataset provided crucial insights into inhibitor mechanisms and their potential modulation by FVIII's B-domain or structural modifications. [27,30,36,68-72,147-150]

2.6.2.B Surface area accessibility studies: FVIII Interacting partners

- **Interacting partners:** The interactions of FVIII with key partners such as vWF and thrombin were meticulously analyzed to understand how B-domain and glycosylation impact binding at critical sites. For vWF, the SASA of major surface epitopes involved in binding was calculated between the refined *Gly/FC-FL-FVIII and Gly-BDD-FVIII* models. These epitopes included regions in the a3-acidic peptide - Q1675-V1689, the A3-domain - H1716-F1718, E1756-D1759, Q1764-L1766, P1884-H1886, the C1-domain - S2075-N2079, P2086-F2087, T2139, and the C2-domain- P2245-Q2285. These binding regions mediate its non-covalent association of FVIII-vWF, essential for the stability and function of FVIII in circulation and protect FVIII from premature degradation. This comparative analysis revealed structural and functional differences between the glycosylated and non-glycosylated forms, aiding in understanding the molecular mechanisms behind vWF interactions. [44,46,47,151] For thrombin, a similar approach was employed to evaluate SASA across the main thrombin cleavage sites within FVIII. These cleavage sites include R391, which bisects the A1- and A2-domains; R759, located between the A2- and B-domains; and R1708, which cleaves the acidic a3-acidic peptide, facilitating the formation of the FVIIIa heterodimer. The SASA values for these sites were compared between *Gly/FC-FL-FVIII and Gly-BDD-FVIII* models to assess how glycosylation influences the exposure and susceptibility of these critical regions to thrombin-mediated cleavage. Additionally, the SASA of thrombin binding regions within the A1-domain E340-E351 and A3-domain E1680-D1684 was calculated to evaluate their accessibility in the presence and absence of B-domain and glycosylations. [29,38,39,152] These detailed analyses provided comprehensive insights into the structural and functional dynamics of FVIII interactions with its key partners. Understanding the effects of glycosylation on the binding and cleavage patterns of vWF and thrombin enhances our knowledge of FVIII biology and informs strategies for therapeutic development and optimization.

2.6.2.C Surface area accessibility studies: FVIII mutations

- **Mutational Studies:** Our functional characterization studies also included analyzing the three predominant mutation types associated with Hemophilia-A - nonsense, missense, and frameshift mutations. The mutations were systematically curated from reliable sources, including the CHAMP database maintained by the CDC (https://www.cdc.gov/ncbddd/hemophilia/champs.html), the European Association for Haemophilia and Allied Disorders (EAHAD) F8 database (https://f8-db-eahad.org), and relevant peer-reviewed literature. [14,51-57,129,149,153-158] These mutations were further categorized based on their type, clinical severity, and association with inhibitor development, allowing for a filtered selection of datasets suitable for this study. To understand the structural and functional implications of these mutations, the

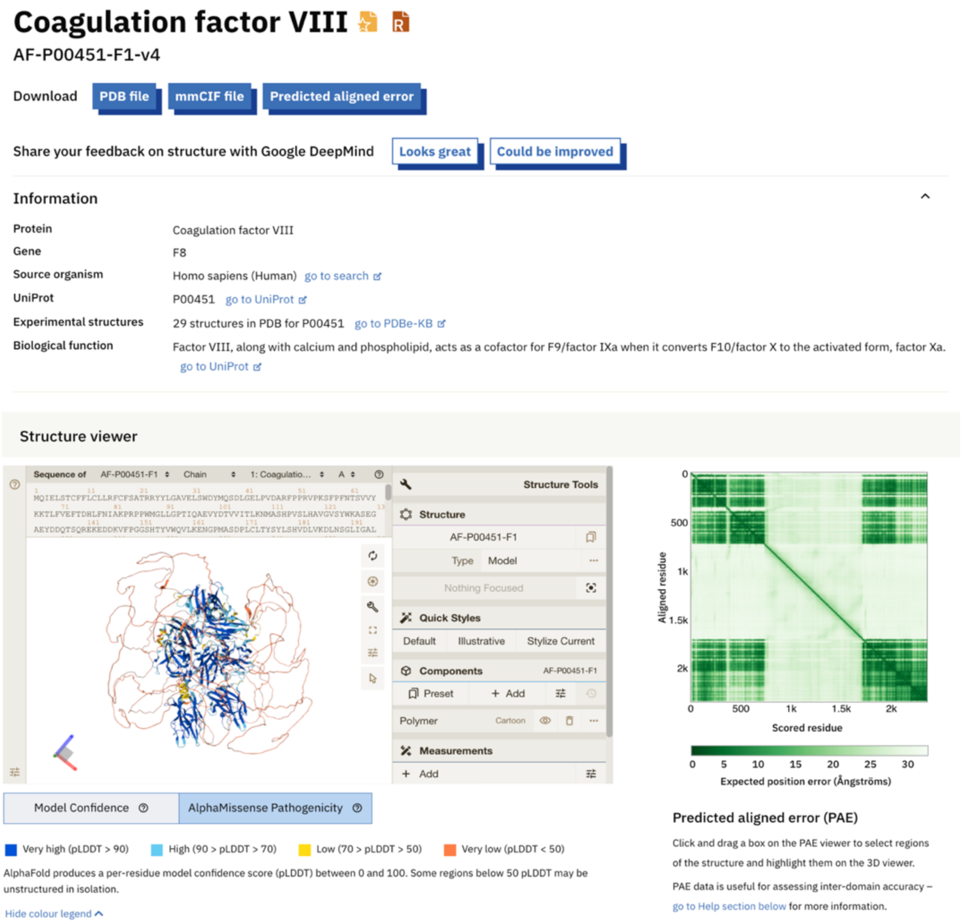
SASA differences of the mutated residues between the refined *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII* models were computed. Additionally, the change in Gibbs free energy (ΔΔG) and disruptiveness/non-disruptiveness factors were calculated using the SAAMBE-3D web server. [159,160] This analysis helped quantify the stabilizing/destabilizing effects of mutations, revealing how alterations in the B-domain and glycosylations could affect them. This analysis offered clinical insights and a detailed view of the glycosylations and B-domain's role in modulating the structural consequences of mutations and their contribution to the pathogenic mechanisms underlying Hemophilia-A and inhibitor formation.

CHAPTER 3 - RESULTS

3.1 Results - Computational studies

3.1.1 Computational studies - Models of FL-FVIII, BDD-FVIII and FL-FVIII-vWF complex structures

3.1.1.A Model building FL-FVIII, BDD-FVIII, and FL-FVIII-vWF complex computational models



Coagulation factor VIII

AF-P00451-F1-v4

Share your feedback on structure with Google DeepMind

Information

Protein	Coagulation factor VIII
Gene	F8
Source organism	Homo sapiens (Human) go to search
UniProt	P00451 go to UniProt
Experimental structures	29 structures in PDB for P00451 go to PDBe-KB
Biological function	Factor VIII, along with calcium and phospholipid, acts as a cofactor for F9/factor IXa when it converts F10/factor X to the activated form, factor Xa. go to UniProt

Structure viewer

🟦 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.

Predicted aligned error (PAE)

Click and drag a box on the PAE viewer to select regions of the structure and highlight them on the 3D viewer.

PAE data is useful for assessing inter-domain accuracy – go to Help section below for more information.

Figure 17. AlphaFold Predicted Structure of FVIII. The figure displays the AlphaFold-predicted 3D structural model of coagulation factor VIII (UniProt ID: P00451). The left panel shows the predicted structural model, with domain annotations and predicted local confidence levels indicated by the pLDDT score (color-coded from blue to red, where blue represents high confidence and red indicates regions with lower confidence). The structure viewer highlights the model's conformational arrangement, including domains critical for function and binding interactions. The right panel displays the PAE plot, high-confidence regions are marked with lower PAE values, while regions with greater flexibility or uncertainty show higher PAE values.

The structural model of FL-FVIII obtained from the AF2 protein structure database - AF-FL-FVIII provided a high-confidence prediction for the ordered domains of the protein. However, the B-domain exhibited lower confidence (mentioned in detail below) in its structural prediction, highlighting uncertainties in this region's conformation. The model displayed the canonical FVIII domains, including the HC $_{\text{(A1-A2-domains)}}$ had very high pLDDT confidence scores > 90 and high scores of 90 > pLDDT > 70 specifically in the a1-a2-acidic rich regions. A similar pattern was also observed for the LC $_{\text{(A3-C1-C2-domains)}}$ with very high and high confidence scores. The predicted structure reflected a modular organization, with well-resolved folding patterns in the HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$. The beta-barrel folds in the HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$ and the inter-domain connectivity appeared highly consistent with previous experimental structural studies, [121] further validating the model's accuracy. The B-domain, known for its flexibility and heavily glycosylated nature, exhibited higher variability in the PAE map and lower pLDDT confidence scores with a mixture of low scores of 70 > pLDDT > 50 and very low scores of pLDDT < 50, reflecting its intrinsic disorderness. The PAE map highlighted regions of high-confidence alignment, particularly within the HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$ with alignment errors predominantly localized to the B-domain. This AF-FL-FVIII model served as a robust foundation for implementing additional modifications, including glycosylations, furin cleavage at R1667, and B-domain deletion, enabling us to assess the impact of these alterations (Figure 17).

- **Effect of Furin cleavage:** Using the Glycam server, high-mannose and hybrid N-glycans specific 19 sites were successfully modeled ensuring their compatibility with the structure. The chosen N-linked glycans created a β-1,4 bond between the N-acetylglucosamine (GlcNAc) base and the nitrogen atom of the (N) residues. Further, EM of the glycosylated structure yielded a Gly-FL-FVIII model with minimal steric clashes and stabilized the structure for further modifications and MD simulation studies. Additionally, the AF-FL-FVIII model without glycosylation - NGly-FL-FVIII was retained. Cleavage of the peptide bond between R1667 and E1668 in both Gly-FL-FVIII and NGly-FL-FVIII models successfully generated two peptide fragments with new N- and C-termini. EM resolved steric clashes and ensured the structural stability of the cleaved models. The conformational adjustment following cleavage was evident, as observed in the relaxed structures post-EM. The process yielded glycosylated-furin cleaved (Gly/FC-FL-FVIII) and non-glycosylated-furin cleaved (NGly/FC-FL-FVIII) structures, both of which were energetically stable and ready for further analysis.
- **B-domain deletion to generate BDD-FVIII:** The glycosylated BDD-FVIII (Gly-BDD-FVIII) model was successfully generated by removing residues R759 to R1667 corresponding to the B-domain from the Gly/FC-FL-FVIII structure. The four glycosylation sites N60, N258, N1829, N2137, containing N-glycans within the A1-, A2-, A3-, and C1-domains, were prominent for the overall structural integrity. EM

resolved any steric clashes or unfavorable interactions introduced during the deletion, resulting in a relaxed and energetically stable structure suitable for further analysis.

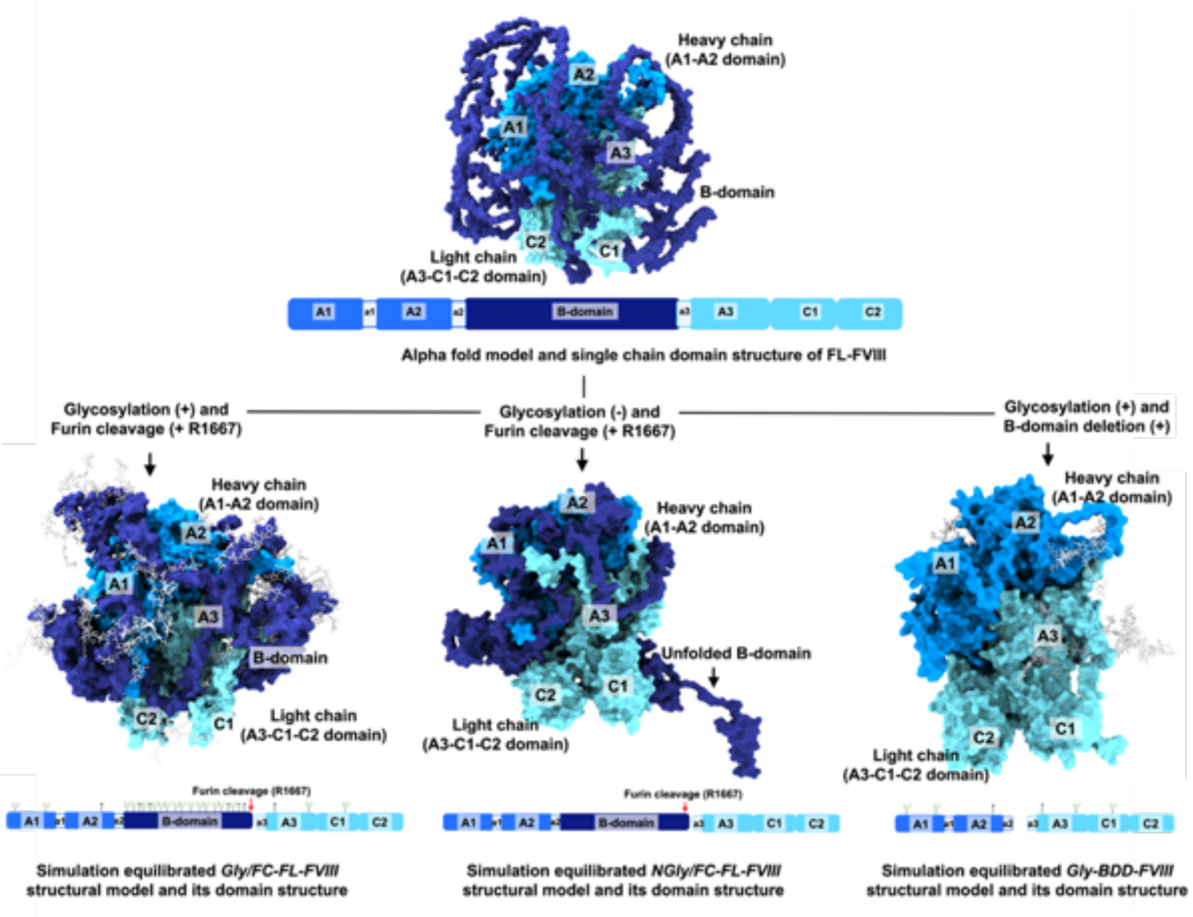


Figure 18. MD simulated and refined computational models developed. The figure illustrates the MD simulated and refined Gly/FC-FL-FVIII, Gly-BDD-FVIII, and NGly/FC-FL-FVIII models and its domain organization derived from AlphaFold-predicted structure of FL-FVIII. Structural changes due to glycosylation, furin cleavage (R1667), and B-domain deletion are shown for three variants: Gly/FC-FL-FVIII (glycosylated and furin cleaved form), Gly-BDD-FVIII (glycosylated B-domain deleted form) and NGly/FC-FL-FVIII (non-glycosylated and furin cleaved form).

3.1.1.B Molecular Dynamic simulation and energy minimization of the FL-FVII, BDD-FVIII, and FL-FVIII-vWF complex models

- **MD and post-MD simulation analysis of Gly/FC-FL-FVIII model:** The Gly/FC-FL-FVIII model reached equilibrium after 30 ns during MD simulation. The production phase extended up to 200 ns to achieve a stable equilibrated model. Post-simulation analysis revealed significant differences in solute behavior such as RMSD stabilizing around 20 Å and exhibiting minimal fluctuations, ranging from 0.5 - 1 Å until the end of the simulation. The RMSF analysis further revealed the Gly/FC-FL-FVIII model showing minimal fluctuations per residue of min: 2 Å and max: 20 Å. The Gly/FC-FL-FVIII model exhibited Rg of 59 Å at the onset of the simulation, followed by a consistent decrease, with fluctuations limited to 0.1 to 0.5 Å. Hydrogen-bonding analysis revealed that the Gly/FC-FL-FVIII model maintained an average of 1950 H-bonds, indicating robust intramolecular interactions due to glycosylations. The total potential energy of the system during the simulation was found to be more consistent at -218,00000 kJ/mol with minimal fluctuations up to 200 ns. Visually, the Gly/FC-FL-FVIII model equilibrated to a compact structure, with the B-domain wrapping around the other domains of FVIII (Figure 18 and Figure 19).
- **MD and post-MD simulation analysis of NGly/FC-FL-FVIII model:** The NGly/FC-FL-FVIII model reached the equilibration phase at 35-40 ns and continued to the production phase after 45 ns up to 350 ns. The NGly/FC-FL-FVIII model did not fully reach equilibration, even after 350 ns of simulation. Post-simulation analysis revealed RMSD stabilizing around 30 Å and exhibited persistent fluctuations of up to 4-5 Å suggesting that the absence of glycosylation might lead to reduced structural stability. Compared to the Gly/FC-FL-FVIII model, the NGly/FC-FL-FVIII model exhibited significant fluctuations in RMSF, reflecting increased variability in atomic positions. Additionally, the NGly/FC-FL-FVIII model showed substantial Rg volatility of 1 to 5 Å throughout the simulation, reinforcing the notion of instability in the absence of glycosylations. Additionally, NGly/FC-FL-FVIII model exhibited a peak hydrogen bond count of only 1550, i.e., 400 fewer H-bonds than Gly/FC-FL-FVIII further underscored the destabilizing effect of the lack of glycosylation, The total potential energy of the system during the simulation was found to be highly fluctuating with min: -955,6000 kJ/mol and max: -9546000 kJ/mol. Visually, the NGly/FC-FL-FVIII model exhibited an unfolded tail-like appendage of the B-domain highlighting a significant structural divergence driven by glycosylation (Figure 18 and Figure 19).
- **MD and post-MD simulation analysis of Gly-BDD-FVIII model:** The Gly-BDD-FVIII model reached the equilibration phase at 25-30 ns and continued to the production phase after 30 ns up to 150 ns. Post-simulation analysis revealed RMSD stabilizing around 30 Å and exhibited minimal fluctuations of up to 2 Å difference. The Gly-BDD-FVIII model exhibited minimal fluctuations in RMSF. Additionally, Rg fluctuations were 9 Å at the onset of the simulation, followed by a consistent

decrease, with fluctuations limited to 0.1 to 0.5 Å. Of 1 to 5 Å throughout the simulation. Additionally, the Gly-BDD-FVIII model exhibited a peak hydrogen bond count of only 1280, due to the absence of the B-domain and its glycosylations. The total potential energy of the system during the simulation was found to be -614,0000 kJ/mol. Visually, the Gly-BDD-FVIII model exhibited a stable structure aligning with the other crystal structures deposited in the RCSB-PDB database (Figure 18).

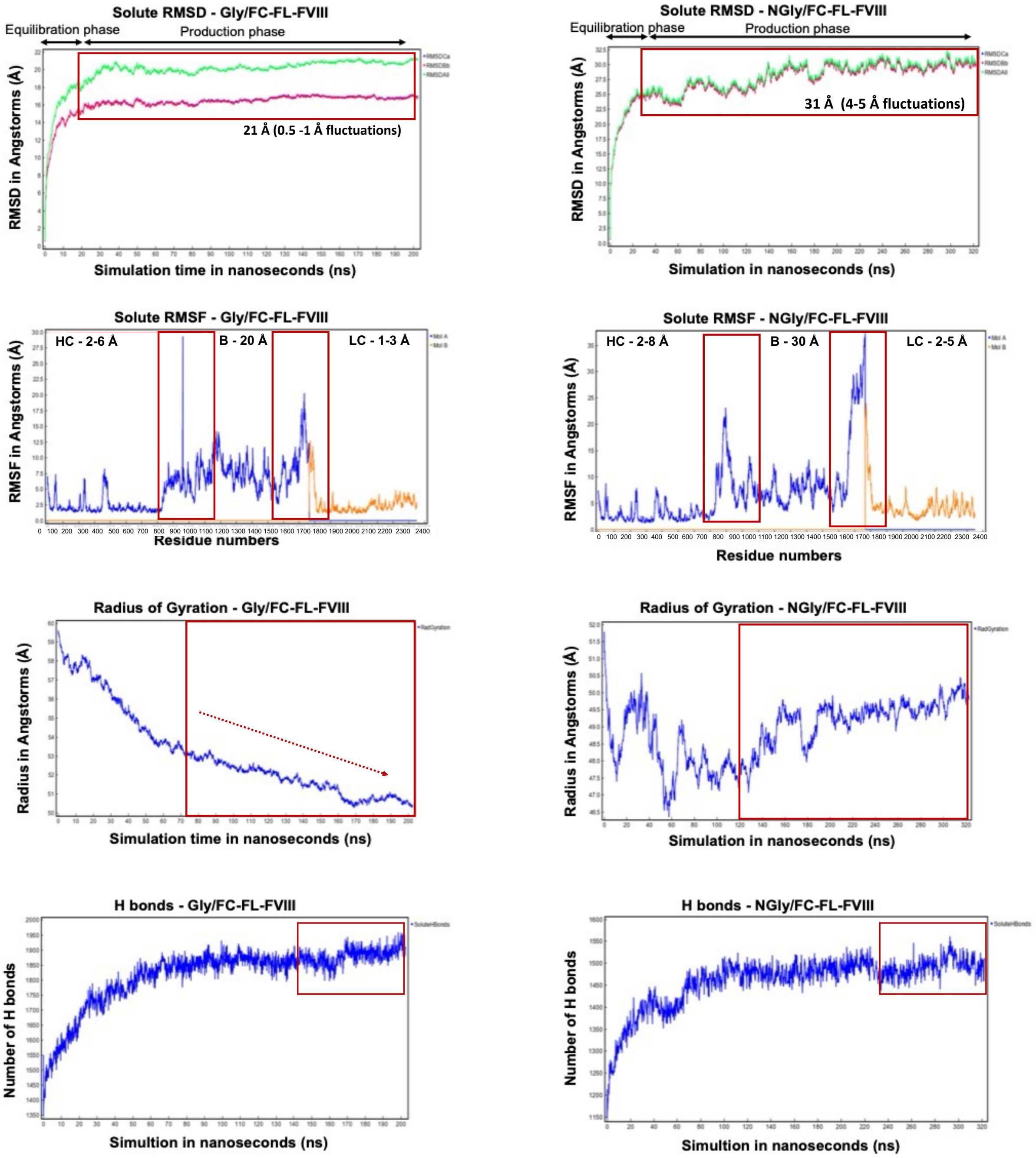


Figure 19. Post simulation analysis of Gly/FC-FL-FVIII, and NGly/FC-FL-FVIII models

3.1.1.C Structural characteristics of validated and refined *Gly/FC-FL-FVIII* model

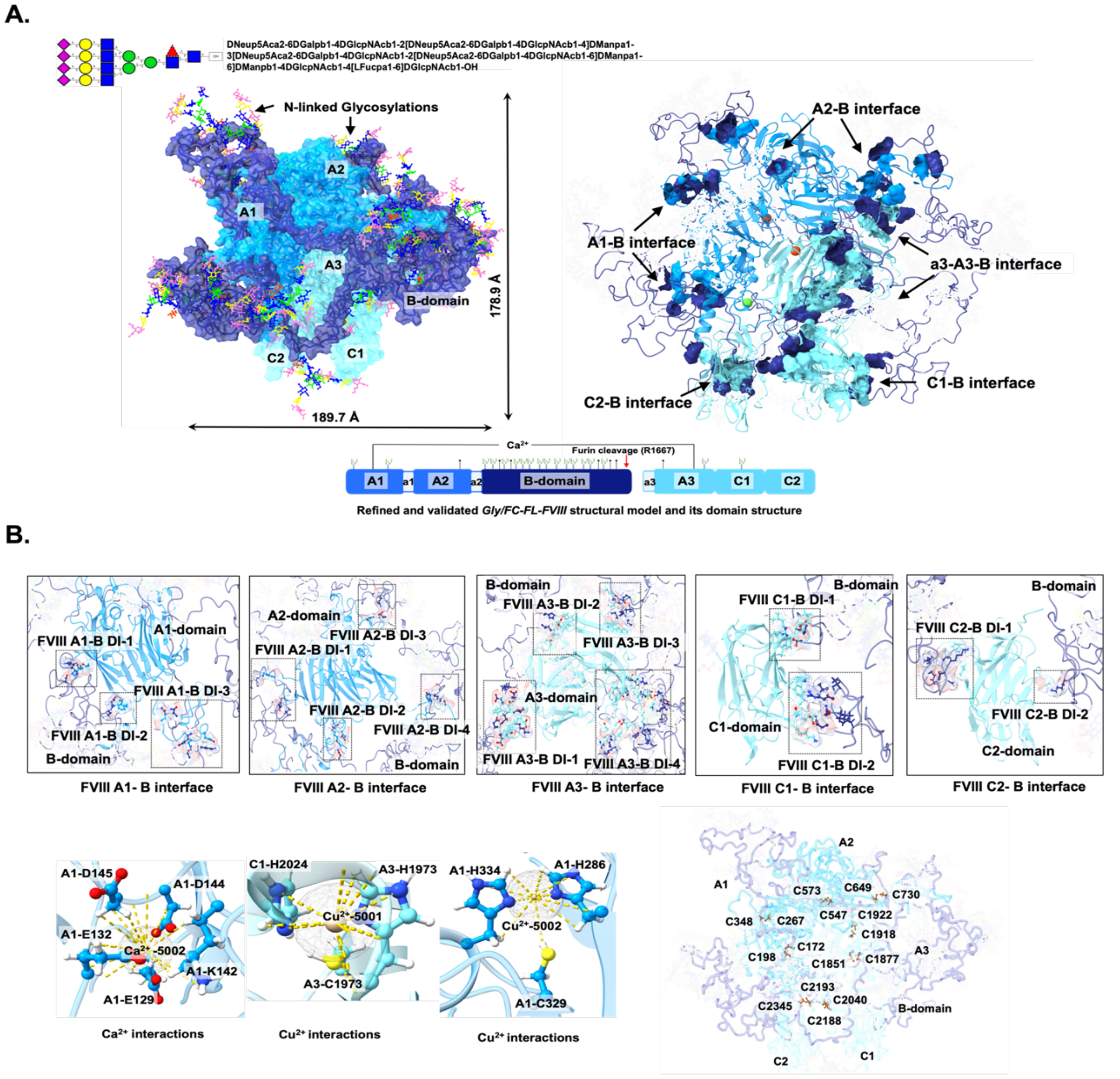


Figure 20. Structural model and characteristics of B-domain within the FL-FVIII. A. Simulation equilibrated Glycosylated/Furin cleaved FL-FVIII (*Gly/FC-FL-FVIII*) model generated from predicted AlphaFold2 model of FL-FVIII showing B-domain (in dense blue). The enlarged image shows the surface representation of the full structure of the FL-FVIII model attached to complex types of N-linked glycosylations. The image to the right shows the interfaces formed between B-domain with HC $_{\text{(A1-A2 domains)}}$ and LC $_{\text{(A3-C1-C2 domains)}}$. **B.** Close-up image of the residues involved in the domain interfaces (DI) and forming H-bonds (<3.5 Å) between A1-B, A2-B, A3-B, C1-B and C2-B-domains. The image to the lower right shows 8 disulfide bonds that are retained as in the B-domain deleted FVIII (PDB-ID:3CDZ) and the lower left image depicts one Calcium, and two copper ions incorporated in the FL-FVIII model and their respective interacting residues.

The model reveals an extended molecular dimension of X: ~ 189.7 Å and Y: ~ 111.7 Å, reflective of its structural integrity and functional conformation. The *Gly/FC-FL-FVIII* model maintained the pseudo-three-fold symmetry of the A1-, A2-, and A3-domains and the co-planar orientation of the C1- and C2-domains. The B-domain interacted specifically with each domain in HC and LC. The HC $_{\text{(A1-A2-domains)}}$-B-domain interface had 29 primary H-bonds, while LC $_{\text{(A3-C1-C2-domains)}}$-B-domain had 11 primary H-bonds as mentioned in Figure 20 and Table 4.

FL-FVIII-Heavy chain (A1-A2 domain)	**Dist. [Å]**	**FL-FVIII-B-domain**
H:LEU 354[N]	2.93	B:ALA1008[O]
H:LYS 512[HZ1]	1.85	B:GLY 924[O]
H:PHE 18[O]	2.24	B:ARG1191[HH22]
H:SER 19[O]	1.87	B:ARG1191[HH12]
H:SER 19[O]	2.48	B:ARG1191[HH22]
H:ASP 75[OD1]	2.09	B:ARG1130[HE]
H:ASP 75[OD2]	2.22	B:ARG1130[HH21]
H:GLN 353[OE1]	1.99	B:ASN1007[HD22]
H:ASN 359[O]	1.76	B:LYS1011[HZ2]
H:GLU 363[OE2]	1.70	B:LYS1018[HZ1]
H:SER 392[O]	2.03	B:ASN 879[HD21]
H:LYS 395[O]	2.06	B:ASN 879[HD22]
H:ALA 420[O]	3.44	B:THR1313[N]
H:PRO 421[O]	3.03	B:THR1313[N]
H:ASP 423[OD2]	3.00	B:THR1314[N]
H:ASP 423[OD2]	2.18	B:GLN1323[HE22]
H:TYR 506[O]	3.01	B:LEU 904[N]
H:SER 507[OG]	2.39	B:ASN 903[HD21]
H:TYR 738[OH]	3.15	B:LYS 787[N]
H:ASN 754[OD1]	2.83	B:PHE 775[N]
H:ARG 759[O]	3.54	B:PHE 761[N]
FL-FVIII-Heavy chain (A1-A2 domain)	**Dist. [Å]**	**FL-FVIII-Light chain (A3-C1-C2 domain)**
H:ASN 703[OD1]	3.10	L:TYR1811[N]
H:LEU 706[O]	2.97	L:ARG1822[N]
H:PRO 704[O]	1.97	L:ARG1822[HE]
H:TYR 683[O]	1.85	L:LYS1852[HZ3]
H:PRO 283[O]	1.92	L:ASN1969[HD22]
H:GLU 684[OE1]	1.55	L:LYS1986[HZ3]
H:SER 168[OG]	2.77	L:GLU1989[N]
H:ASP 169[OD2]	3.16	L:LYS1991[N]
H:GLU 162[OE2]	1.43	L:LYS1991[HZ2]
H:ILE 310[O]	1.92	L:ASN1996[HD22]
H:SER 123[OG]	1.88	L:LYS2011[HZ3]
H:TYR 124[O]	1.86	L:LYS2011[HZ1]
H:GLY 280[O]	3.81	L:GLY2022[N]
H:THR 281[O]	2.99	L:GLU2023[N]
H:ASP 135[O]	1.77	L:ASN2191[HD21]

H:GLU 141[OE2]	1.95	L:LYS2258[HZ1]
H:SER 138[OG]	2.01	L:LYS2258[HZ2]
H:THR 154[OG1]	2.96	L:TYR2351[N]
H:ASN 703[HD21]	2.06	L:ILE1809[O]
H:SER 729[N]	3.83	L:GLU1813[OE1]
H:ASN 299[HD22]	2.00	L:GLU1988[OE1]
H:SER 168[N]	3.30	L:GLU1989[O]
H:SER 308[OG]	3.13	L:TYR1998[OH]
H:LYS 126[HZ3]	2.21	L:ILE2014[O]
H:TYR 133[HH]	1.78	L:ILE2014[O]
FL-FVIII-Light chain (A3-C1-C2 domain)	**Dist. [Å]**	**FL-FVIII-B-domain**
L:GLN1678[HE22]	2.02	B:GLN 1209[O]
L:GLN1678[HE21]	2.24	B:GLU 1210[O]
L:LYS2298[HZ2]	2.49	B:GLY 1509[O]
L:GLN1678[OE1]	3.12	B:LYS 1212[N]
L:GLN1678[O]	2.96	B:ILE 1213[N]
L:GLN1815[OE1]	1.81	B:GLN 1302[HE21]
L:GLU1820[OE2]	2.77	B:GLN 1305[N]
L:GLN1925[OE1]	2.80	B:ARG 1338[N]
L:ASN1923[O]	2.07	B:ARG 1338[HE]
L:ILE1924[O]	1.89	B:ARG 1338[HH21]
L:GLN1925[OE1]	2.91	B:LEU 1339[N]

Table 4. Structural characteristics *Gly/FC-FL-FVIII* model - H-bonds and Interface residues

- The metal ions (one Ca^{2+} and two Cu^{2+}) had similar coordination residues, Ca2+ H-bonds with K126<->E129<->E132<->D144<->D145. One of the Cu^{2+} bonded with H286<->C329<->H334 and the other Cu^{2+} bonded with H1973<->C2019<->H2024.
- The eight disulphide bonds: C172<->C198, C267<->C348, C547<->C573, C649<->C730, C1851<->C1877, C1918<->C1922, C2040<->C2188, C2193<->C2345.
- Six tyrosine sulfation sites : Y365, Y737, Y738, Y742, Y1683, Y1699

The eight disulfide bonds six tyrosine sulfation sites and metal binding coordinates mentioned and further detailed in Figure 20 and Table 5 remained intact regardless of the B-domain and were consistent with previous BDD-FVIII crystal structures.

Ions	Interacting residues in FL-FVIII
Ca^{2+}	K126 / E129 / E132 / D144 / D145
Cu^{2+}	H286 / C329 / H334
Cu^{2+}	H1973 / C2019 / H2024

Disulphide bonds	Cys172 (C172) : Cys198 (C198) /Cys267 (C267) : Cys348 (C348) /Cys547 (C547) : Cys573 (C573) /Cys649 (C649) : Cys730 (C730)/Cys1851 (C1851) : Cys1877 (C1877) /Cys1918 (C1918) : Cys1922 (C1922) /Cys2040 (C2040) : Cys2188 (C2188) /Cys2193 (C2193) : Cys2345 (C2345)
Sulfation	Tyr365 (Y365)/Tyr737 (Y737)/Tyr738 (Y738)/Tyr742 (Y742)/Tyr1683 (Y1683)/Tyr1699 (Y1699)

Table 5. Structural characteristics *Gly/FC-FL-FVIII* model - Disulphides, Sulfation sites, and Metal ions

- The A1-B-domain included 3 primary domain interfaces (DI) which included residues from signal peptide, and the first thrombin cleavage site (R391). The A2-B-domain included 4 DIs including the second thrombin cleavage site (R759). The a3-A3-B-domain interfaces had 4 regions with the third thrombin cleavage site (R1708). The C1-B-domain had 2 interface regions, and the C2-B-domain had 2 interface regions as mentioned in Figure 20 and Table 6.

Interacting domains with B-domain	Interacting residues in FL-FVIII
A1	F18, S19, T74, D75, H76, Q353, N359, E363, R1191, N1192, R1130, S1144, N1007, K1011, S1015, K1018
A2	S392, K395, D423, Q429, E475, S507, K512, E576, Y738, S741, N879, T1313, Q1323, N903, G924, K787
A3	Q1678, E1679, E1680, K1693, I1681, A1920, P1921, C1922, Q1925, E1927, D1928, E1820, F1825, K1823, P1929, G1744, R1919, Q1742, A1741, S1731, L1737, S1732, R1738, E1787, V1786, T1879, I1681, E1680, E1679, K1693, E1915, Q1815, N1923, I1924, Q1209, E1210, Q1302, R1338, S1405, P1408, L1407, P1406, V1307, L1339, Q1305, I1306, L1341, F1337, P1420, I1418, R1419, P1420, L1467, L1469, E1470, H1234, I1233, E1215, I1213, E1216, K1212
C1	D2127, L2126, R2166, Y2167, Q2061, K2084, E2085, F2087, W2089, Q1455, S1450, Q1636, N1637, R1650, V1643
C2	N2296, G2297, K2298, F2294, V2299, A2207, L1513, E1512, T1507, S1508, K1506, K1178

Table 6. Structural characteristics *Gly/FC-FL-FVIII* model - B-domain interacting residues.

3.1.1.D Computational studies - Structural characteristics of validated and refined FL-FVIII-VWF complex model

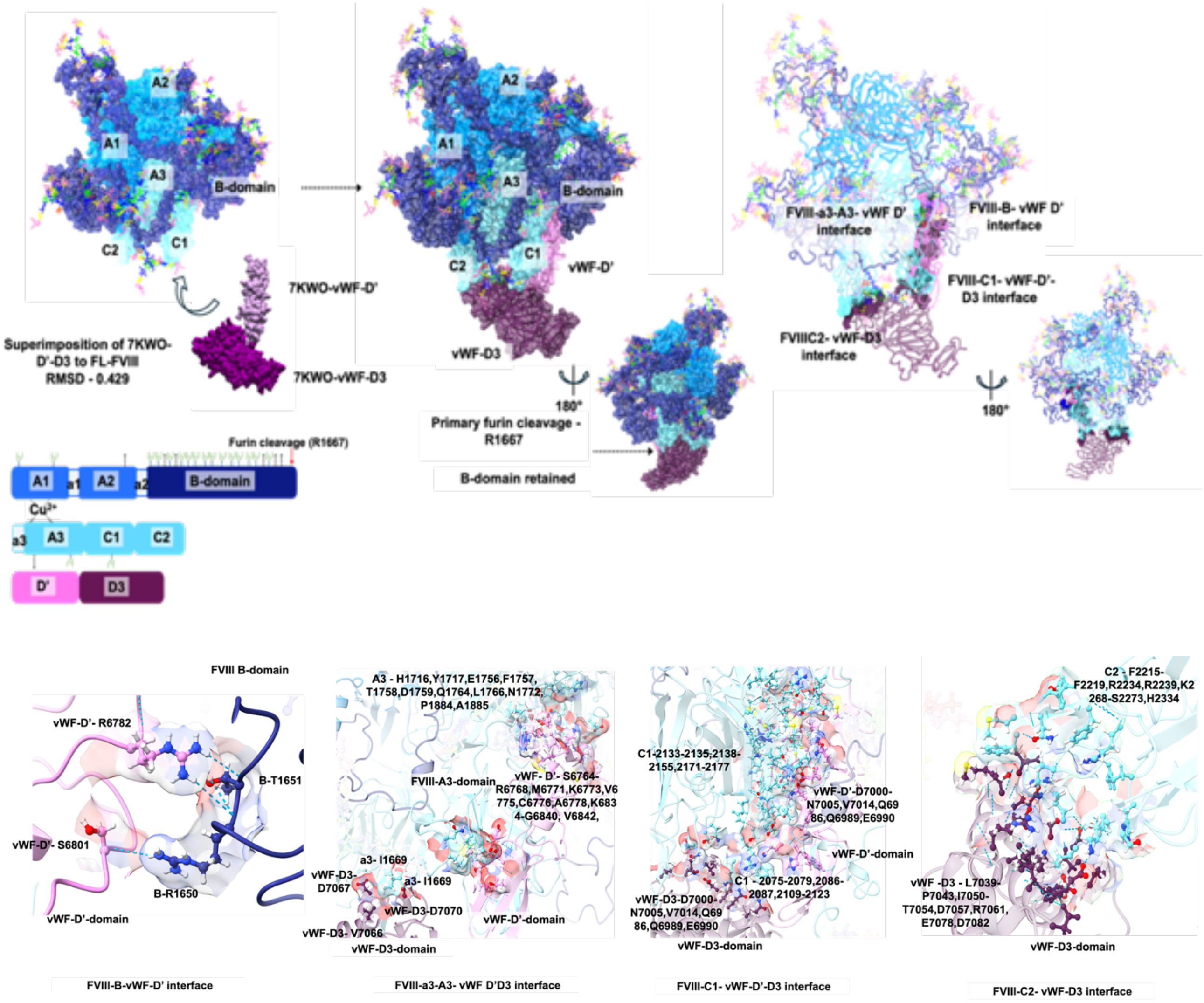
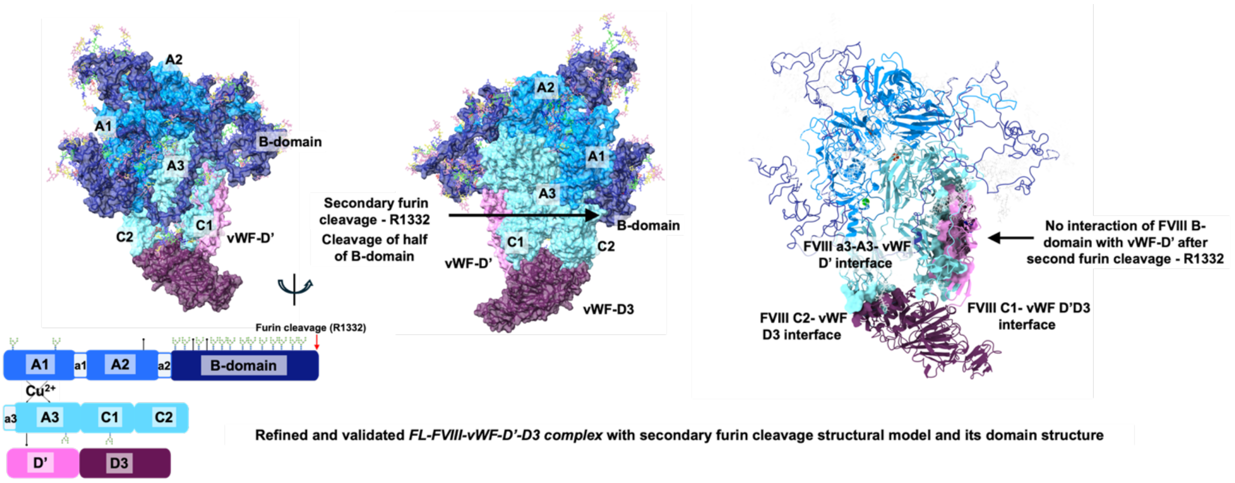


Figure 21. Structural model and characteristics of FL-FVIII-vWF-D′-D3 complex. Top left corner images depict the superimposition of the D'-D3 domains of vWF from Efanesoctcog alfa (PDBID: 7KWO) onto simulation equilibrated glycosylated/furin cleaved (*Gly/FC-FL-FVIII*) generating a *Gly/FC-FL-FVIII-vWF-D'-D3* complex model, along with its domain structure. The center image shows the surface representation of the *Gly/FC-FL-FVIII-vWF-D'-D3* complex model. The top right image shows the interfaces formed between FVIII B-vWF-D' domains, FVIII a3-A3-vWF-D3 domains, FVIII C1-vWF-D'D3 domains, and FVIII C2-vWF-D3 domains. The close-up image depicts the residues involved in the interfaces and forming H-bonds (<3.5Å) between FVIII B-vWF-D' domains.

The vWF D'-D3 complex was successfully superimposed using the D'-D3 domains from the BDD-FVIII-vWF cryo-EM structure (PDB-ID: 7KWO) as the reference and the simulation-equilibrated *Gly/FC-FL-FVIII* model. This superimposition involved precise adjustments and rotations of the C2- and C1-domains to achieve optimal alignment.

During this process, care was taken to ensure that key secondary structures and functional regions were preserved in their relative positions while aligning the overall architecture of the complex. The quality of the alignment was assessed by calculating the RMSD value for the aligned region, which was determined to be 0.429 Å, indicative of an excellent alignment and reflects a high degree of structural congruence. Such a low RMSD value suggests that the modeled complex closely matched the conformation of the reference structure, particularly in the D'-D3-domains and their interactions with other domains of *Gly/FC-FL-FVIII*. In addition to ensuring the proper orientation of the C2- and C1-domains, the alignment process revealed key interaction interfaces that were maintained across the superimposed structures. These included critical regions involved in inter-domain interactions such as B-domain, a3-acidic, A3-domain, C1-domain, and C2-domain of *Gly/FC-FL-FVIII*. Notably, the B-domain formed three key hydrogen bonds with the D'-domain: R1650 <-> S6801, R1650 <-> R6782, and T1651 <-> R6782, which were novel interactions identified. The interface between the B-domain and the D'-domain involved eight residues from the B-domain and six residues from the D'-domain, with solvent-accessible surface areas of 189.8 Å^2 and 234.7 Å^2, respectively. Similar interactions were observed for the a3 acidic, A3-, C1-, and C2-domains, where a total of 11 hydrogen bonds and six salt bridges were formed with the D'-D3 domains. Overall, 66 interface residues were identified between the light chain of FL-FVIII and vWF, with solvent-accessible surface areas of 2335.3 Å^2 for FL-FVIII and 2270.7 Å^2 for vWF (Figure 21, Table 7, Table 8). Beyond these specific interactions, the remaining interface characteristics were found to be similar to those observed in the *Gly/FC-FL-FVIII* model. The refined and validated superimposed model as mentioned earlier, generated the *Gly/FC-FL-FVIII-D'-D3* complex model. This model retained the primary furin cleavage site at R1667 from the *Gly/FC-FL-FVIII* structure, facilitated the proper orientation and interaction of the B-domain with the D'-domain observed in the superimposed *Gly/FC-FL-FVIII-D'-D3* complex model. In contrast, an analysis of the primary AF-FL-FVIII model without furin cleavage revealed significant structural tension and instability, particularly in regions surrounding the B-domain. These findings suggest that furin cleavage at R1667 is essential for alleviating structural tension and stabilizing the FVIII structure. Further analysis demonstrated that these critical interactions between the B-domain and the D'-domain were absent in the *Gly/FC-FL-FVIII-D'-D3* complex model containing an additional furin cleavage at R1332. This secondary cleavage resulted in the removal of approximately half of the B-domain, leading to a significant reduction in potential interaction sites. Moreover, the presence of furin cleavage at R1332 could be associated with substantial conformational changes in the B-domain, particularly its unwrapping around the FVIII core during Golgi processing (Figure 22, Table 7, Table 8).



Refined and validated *FL-FVIII-vWF-D'-D3 complex* with secondary furin cleavage structural model and its domain structure

Figure 22. Impact of Furin cleavage in removal of B-domain. FL-FVIII-vWF-D′-D3 complex after the secondary furin cleavage at R1332 along with its respective domain structure is shown on the left. The image on the right shows that no domain interfaces were identified between the B-domain and the vWF-D' domains, suggesting that secondary furin cleavage removes half of the B-domain while preserving the remaining interfaces**.**

FL-FVIII-vWF B-domain	**Dist. [Å]**	**FL-FVIII-vWF-D'-D3-domain**
B:R1650[HH22]	2.07	V:S6801[O]
B:R1650[O]	2.40	V:R6782[HH12]
B:T1651[O]	1.77	V:R6782[HH22]
FL-FVIII-Light chain (A3-C1-C2-domain)	**Dist. [Å]**	**FL-FVIII-vWF-D'-D3 domain**
L:S2077[N]	3.39	V:N7004[OD1]
L:K2111[HZ2]	2.02	V:D7000[O]
L:K2111[HZ3]	2.32	V:Q6989[O]
L:T2141[N]	3.27	V:S6806[O]
L:K2155[HZ3]	2.21	V:E6818[OE1]
L:K2268[HZ2]	1.48	V:E7078[OE2]
L:M1691[SD]	2.14	V:K6834[HZ3]
L:G2076[O]	2.04	V:N7004[HD22]
L:Q2119[OE1]	2.18	V:N6794[HD22]
L:G2147[O]	2.11	V:Q6793[HE21]
L:T2173[O]	2.74	V:M6800[N]

Table 7. Structural characteristics *Gly/FC-FL-FVIII-D'-D3 complex* model - H-bonds and Interface residues

Domains	Interacting residues in FL-FVIII-vWF
D'-D3-B	**FVIII-B- domain:** K1579, L1580, L1584, G1649, R1650, T1651, E1652, L1654 **vWF-D'- domain :** D6779, L6781, R6782, M6800, S6801, M6802
a3-A3-D'	**FVIII-a3-A3-domain :** I1669,L1674,Q1675,E1690,M1691,H1716,Y1717,E1756,F1757,T1758,D1759,Q1764,L1766,N1772,P1884,A1885 **vWF-D'- domain :** S6764, L6765, S6766, R6768, M6771, K6773, V6775, C6776, A6778, K6834, E6835, Y6836, G6839, E6840, T6841, V6842, D7066, D7070, V7067
C1-D'-D3	**FVIII-C1- domain :** S2075, G2076, S2077, N2079, P2086, F2087, R2109, K2111, F2112, Y2116, S2118, Q2119, I2121, M2123, T2133, R2135, S2138, T2139, G2140, T2141, L2142, M2143, V2144, F2146, G2147, N2148, V2149, D2150, S2152, G2153, K2155, H2171, P2172, T2173, H2174, Y2175, S2176, I2177 **vWF-D'- domain:** L6786, E6787, C6788, T6789, T6791, Q6793, N6794, L6797, E6798, C6799, M6800, S6801, M6802, G6803, C6804, V6805, S6806, G6807, C6808, R6816, E6818, N6819, F6830, H6831, Q6832, G6833
C2-D'-D3	**FVIII-C2-domain :** F2215, T2216, N2217, M2218, F2219, R2234, R2239, K2268, L2270, L2271, T2272, S2273, H2334 **vWF-D3-domain :** H6874, L6884, Q6986, Q6989, E6990, D7000, G7001, I7002, N7004, N7005, V7014, L7039, S7041, S7042, P7043, I7050, M7051, Q7053, T7054, D7057, R7061, D7070, E7078, L7081, D7082, I7085, Y7086

Table 8. Structural characteristics *Gly/FC-FL-FVIII-D'-D3 complex* model - B-domain Interacting residues

3.1.1.E Computational studies - Results of Model building, refinement and validations

The refined *Gly/FC-FL-FVIII, Gly-BDD-FVIII,* and *Gly/FC-FL-FVII-vWF-D'-D3* complex structural models were rigorously validated using MolProbity tool as mentioned in Section 2.2.1.D. This comparative validation of the refined models to the primary AF-FL-FVIII structural model evaluated various aspects of stereochemical quality which are mentioned below.

- The primary **AF-FL-FVIII** structural model was validated using MolProbity to assess its quality and suitability as a reference for comparison with the constructed models. Ramachandran plot analysis of the model showed that 16.69 % of residues were outliers, 69.86 % were in the most favored regions, and 13.45 % fell within allowed regions. All-atom contact analysis revealed significant high-energy clashes, with an overall clash score of 1.72, indicating a high level of steric clashes between atoms, surpassing the threshold for a high-quality structure. Further analysis of side-chain conformations revealed that 6.39 % were rotamer outliers, while 7.25 % exhibited Cβ-deviations, suggesting poor side-chain packing and significant deviations from ideal conformations. Peptide plane conformation analysis identified 17.97 % cis-proline content, 7.11 % cis-general content, 16.41 % twisted-proline content, and 17.64 % twisted-general content, indicating notable deviations from expected backbone geometry. Geometry analysis showed RMS bond length deviations of 0.023 Å and bond angle deviations exceeding 3.07°, both of which are well beyond the acceptable range for high-resolution structural models. The final MolProbity score of 2.38 further confirmed the model's low quality, reflecting extensive steric and geometric inaccuracies (Figure 23, Table 9).

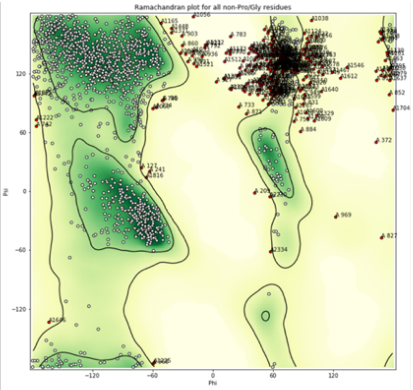
Structural models	**MolProbity score**	**RMS (bonds and angles)**	**Clashscore**	**Ramchandran percentage values**
AF-FVIII	2.38	0.0232 Å and 3.07°	1.69	Outliers - 16.69 % Favored - 69.86 %
Gly/FC-FL-FVIII	1.68	0.0147 Å and 1.83°	0.75	Outliers - 4.49 % Favored - 87.17 %
Gly-BDD-FVIII	1.11	0.0114 Å and 1.69°	0.69	Outliers - 0.83 % Favored - 95.41 %
***Gly/FC-FL-FVII-vWF-D'-D3* complex**	1.67	0.0146 Å and 1.86°	0.97	Outliers - 3.77 % Favored - 88.37 %

Table 9. Validity of the computational refined *Gly/FC-FL-FVIII, Gly-BDD-FVIII,* and *Gly/FC-FL-FVII-vWF-D'-D3* complex models using MolProbity.

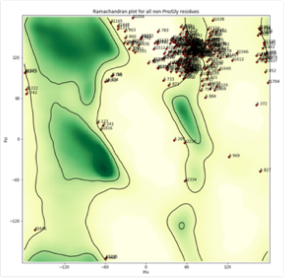
- The refined ***Gly/FC-FL-FVIII*** structural model was drastically improved compared to the AF-FL-FVIII model. Ramachandran plot analysis of the *Gly/FC-FL-FVIII* model indicated 4.48 % of Ramachandran outliers and 87.12 % were in the most favored regions, with 8.4 % in the allowed regions. The φ/ψ torsion angles were predominantly within the favored regions of the Ramachandran space, with only a small proportion of residues in allowed regions, which is typical for models at this resolution. The all-atom contact analysis revealed no high-energy clashes, with an overall energy score consistent with favorable packing and structural stability. An all-atom overall clash score was calculated as 0.64, well below the typical threshold for good-quality structures indicating minimal steric clashes between atoms. Furthermore, the rotamer analysis revealed that 2.65 % of side-chain conformations were outliers, and 2.70 % were Cβ-deviations suggesting excellent side-chain packing and minimal deviations from ideal conformations. Peptide plane conformations indicated the Cis-Proline to be 21.88 %, Cis-general content to be 0.36 %, twisted-proline content to be 9.38 %, and Twisted-general content to be 2.75 %. Geometry analysis showed that RMS bond lengths were within 0.014 Å, and bond angles deviated by less than 1.8°, both of which are well within the typical resolution limits for high-resolution models. The final MolProbity score of 1.65 further corroborated the high-quality, well-validated nature of the *Gly/FC-FL-FVIII* model. These results collectively demonstrated that the *Gly/FC-FL-FVIII* structural model is highly reliable, with minimal steric, and geometric issues (Figure 23, Table 9).

- The refined ***Gly-BDD-FVIII*** structural model was also drastically improved compared to the AF-FL-FVIII model. Ramachandran plot analysis of the *Gly-BDD-FVIII* model indicated 0.83 % of Ramachandran outliers and 95.41 % were in the most favored regions, with 3.75 % in the allowed regions. The φ/ψ torsion angles were predominantly within the favored regions of the Ramachandran space, with only a small proportion of residues in allowed regions, which is typical for models at this resolution. The all-atom contact analysis revealed no high-energy clashes, with an overall energy score consistent with favorable packing and structural stability. An all-atom overall clash score was calculated as 0.69, well below the typical threshold for good-quality structures indicating minimal steric clashes between atoms. Furthermore, the rotamer analysis revealed that 1.24 % of side-chain conformations were outliers, and 0.37 % were Cβ-deviations suggesting excellent side-chain packing and minimal deviations from ideal conformations. Peptide plane conformations indicated the Cis-Proline to be 4.23 %, Cis-general content to be 0.07 %, twisted-proline content to be 1.41 %, and Twisted-general content to be 0.29 %. Geometry analysis showed that RMS bond lengths were within 0.014 Å, and bond angles deviated by less than 1.69°, both of which are well within the typical resolution limits for high-resolution models. The final MolProbity score of 1.11 further corroborated the high-quality, well-validated nature of the *Gly-BDD-FVIII* model. These results collectively demonstrated that the *Gly-BDD-FVIII* structural model is highly reliable, with minimal steric, and geometric issues (Figure 23, Table 9).

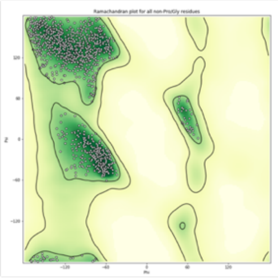
- The refined ***Gly/FC-FL-FVIII-vWF-D'-D3*** complex structural model displayed high quality. Ramachandran plot analysis of *Gly/FC-FL-FVIII-vWF-D'-D3* complex indicated 3.76 % of Ramachandran outliers and 88.37 % were in the most favored regions, with 7.87 % in the allowed regions. The φ/ψ torsion angles were predominantly within the favored regions of the Ramachandran space, with only a small proportion of residues in allowed regions, which is typical for models at this resolution. The all-atom contact analysis revealed no high-energy clashes, with an overall energy score consistent with favorable packing and structural stability. An all-atom overall clash score was calculated as 0.56, well below the typical threshold for good-quality structures indicating minimal steric clashes between atoms. Furthermore, the rotamer analysis revealed that 2.41 % of side-chain conformations were outliers, and 2.25 % were Cβ-deviations suggesting excellent side-chain packing and minimal deviations from ideal conformations. Peptide plane conformations indicated the Cis-Proline to be 18.59 %, Cis-general content to be 0.30 %, twisted-proline content to be 7.69 %, and Twisted-general content to be 2.36 %. Geometry analysis showed that RMS bond lengths were within 0.014 Å, and bond angles deviated by less than 1.80°, both of which are well within the typical resolution limits for high-resolution models. The final MolProbity score of 1.57 further corroborated the high-quality, well-validated nature of the *Gly/FC-FL-FVIII-vWF-D'-D3* complex. These results collectively demonstrated that the *Gly/FC-FL-FVIII-vWF-D'-D3* complex structural model is highly reliable, with minimal steric, and geometric issues (Figure 23, Table 9).



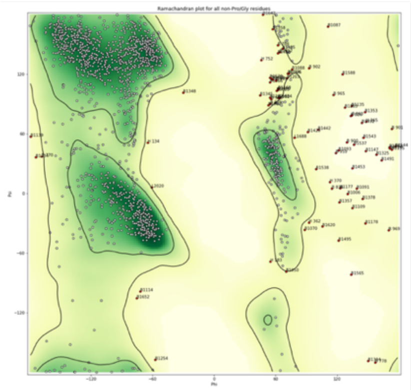
Ramachandran plot for the AF-FL-FVIII model



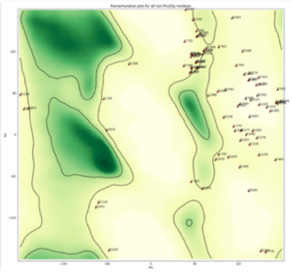
Ramachandran plot for the
AF-FL-FVIII model - outliers



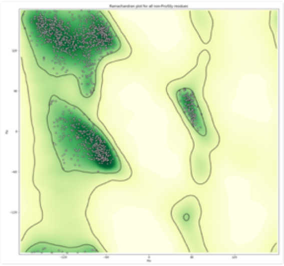
Ramachandran plot for the
AF-FL-FVIII model - favoured



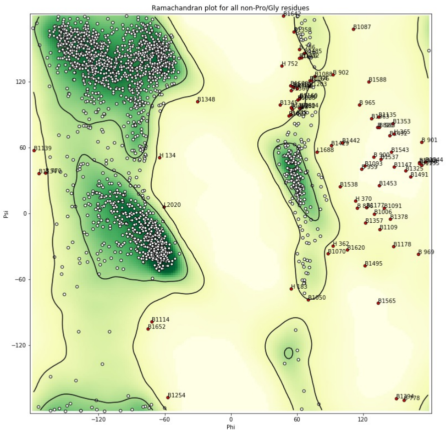
Ramachandran plot for the *Gly/FC-FL-FVIII* model



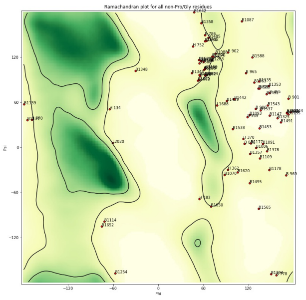
Ramachandran plot for the
Gly/FC-FL-FVIII model - outliers



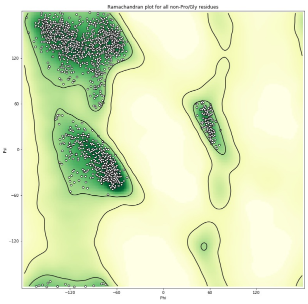
Ramachandran plot for the
Gly/FC-FL-FVIII model - favoured



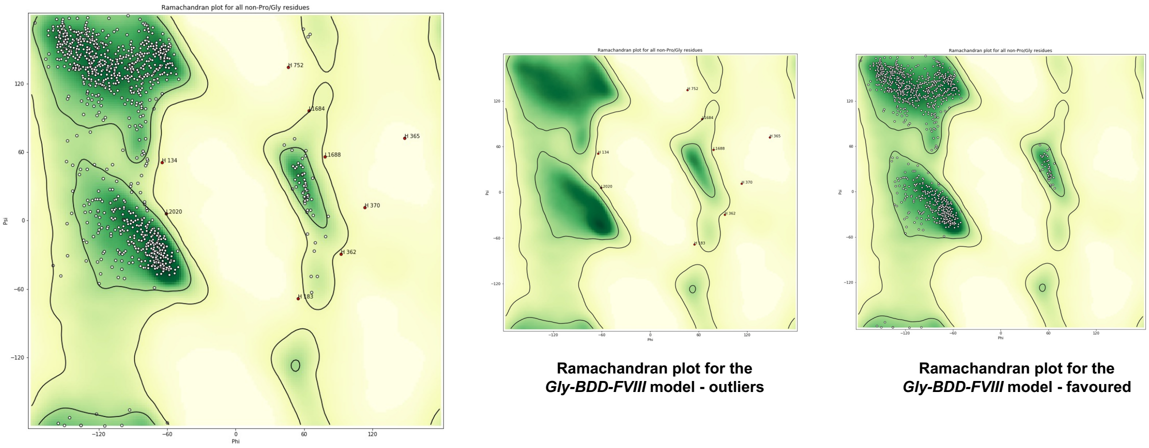
Ramachandran plot for the *Gly/FC-FL-FVIII-vWF-D'-D3* complex model



Ramachandran plot for the *Gly/FC-FL-FVIII-vWF-D'-D3* complex model - outliers



Ramachandran plot for the *Gly/FC-FL-FVIII-vWF-D'-D3* complex model - favoured



Ramachandran plot for the ***Gly-BDD-FVIII*** **model - outliers**

Ramachandran plot for the ***Gly-BDD-FVIII*** **model - favoured**

Ramachandran plot for the ***Gly-BDD-FVIII*** **model**

Figure 23. Model validation of *Gly/FC-FL-FVIII* and *FL-FVIII-vWF-D'-D3* complex models. Ramachandran plots for the A. AF-FL-FVIII B. *Gly/FC-FL-FVIII* and C. *FL-FVIII-vWF-D'-D3* structural models, illustrating the distribution of phi (φ) and psi (ψ) dihedral angles for the backbone conformations. The plots highlight the favored, and outlier regions, providing insights into the steric constraints and overall conformational stability of the protein. The analysis aids in assessing the quality of our *Gly/FC-FL-FVIII* and *FL-FVIII-vWF-D'-D3* models and validates through an improved MolProbity and clash scores of the same than compared to AF-FL-FVIII.

3.2 Results - Protein Biochemistry studies

3.2.1 FVIII protein biochemistry studies - FVIII protein purification strategies

3.2.1.A Purification of rFL-FVIII, and pdFL-FVIII by SEC

SEC purification of rFL-FVIII: The purification of rFL-FVIII was carried out using a Superdex™ 200 Increase 10/300 GL (Cytiva) column based on its size distribution. The chromatogram revealed two distinct peaks, reflecting two different molecular species (Figure 24A).

- **Peak-1:** This peak corresponded to the intact, single-chain FL-FVIII (sc-FL-FVIII). Its early retention volume ~ 9.5 mL was consistent and aligned with the expected MW of rFL-FVIII of ~ 280 kDa, thus eluted as a monodisperse species. The sharp, symmetrical nature of this peak indicated a homogeneous population with minimal aggregation or degradation. The absorbance intensity of Peak-1 corresponding to more than 20 mAU suggested that it constituted the major component of the rFL-FVIII product (Figure 24A).
- **Peak-2:** This smaller peak represented FVIII variants with variable lengths of HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$. Its retention volume was later - ~ 12 mL suggested in variable sizes. The broader profile of this peak indicated heterogeneity within the population, with the presence of multiple high MW species corresponding to HC $_{\text{(A1-A2-domains)}}$ - ~ 100 kDa and lower MW species of LC $_{\text{(A3-C1-C2-domains)}}$ - ~ 75 kDa. This was consistent with the known presence of highly disordered B-domain yielding

variable lengths of HC $_{\text{(A1-A2)}}$ and LC $_{\text{(A3-C1-C2)}}$. The lower absorbance intensity of <5 mAU of Peak-2 suggested that it constituted the minor component of the rFL-FVIII product (Figure 24A).

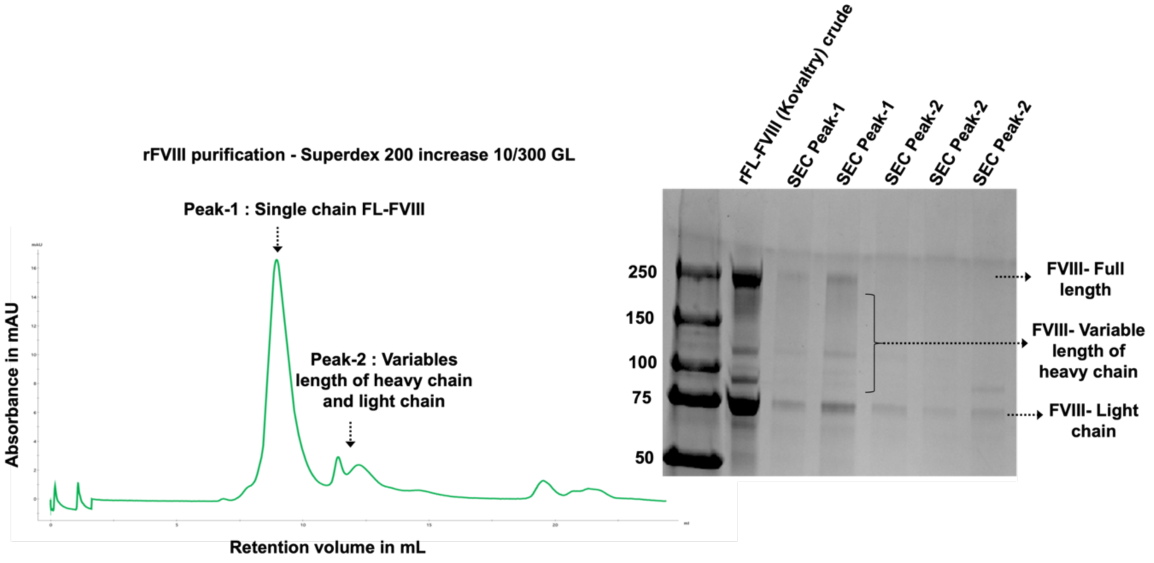


Figure 24A. SEC purification of rFVIII. SEC chromatogram of rFVIII performed using a Superdex 200 Increase 10/300 GL column (Left). The chromatogram shows two distinct peaks: Peak 1 represents single-chain full-length FVIII (FL-FVIII), and Peak 2 corresponds to FVIII fragments with variable lengths of the heavy and light chains. SDS-PAGE analysis of SEC fractions under reducing conditions. Lane 1 shows the molecular weight marker, followed by the rFVIII input sample and SEC fractions (Peaks 1 and 2). The gel indicates full-length FVIII in Peak 1, along with variable heavy and light chain fragments in Peak 2 (Right). Molecular weights are annotated on the left of the gel.

The separation of rFL-FVIII proteins achieved by the Superdex™ 200 Increase 10/300 GL (Cytiva) highlighted the efficiency of SEC in isolating intact FL-FVIII forms from its variable MW/complex forms. SDS-PAGE analysis and coomassie staining of rFL-FVIII, as mentioned in Section 3.3.5, SEC Peak-1 predominantly displayed a distinct band corresponding to sc-FL-FVIII at 250 kDa, confirming its identity as the intact single-chain. In contrast, SEC Peak-2 showed bands representing variable lengths of HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$, of ~ 150-200 kDa and ~ 75 kDa respectively. The unpurified crude rFL-FVIII displayed a mixture of multiple bands (Figure 24A).

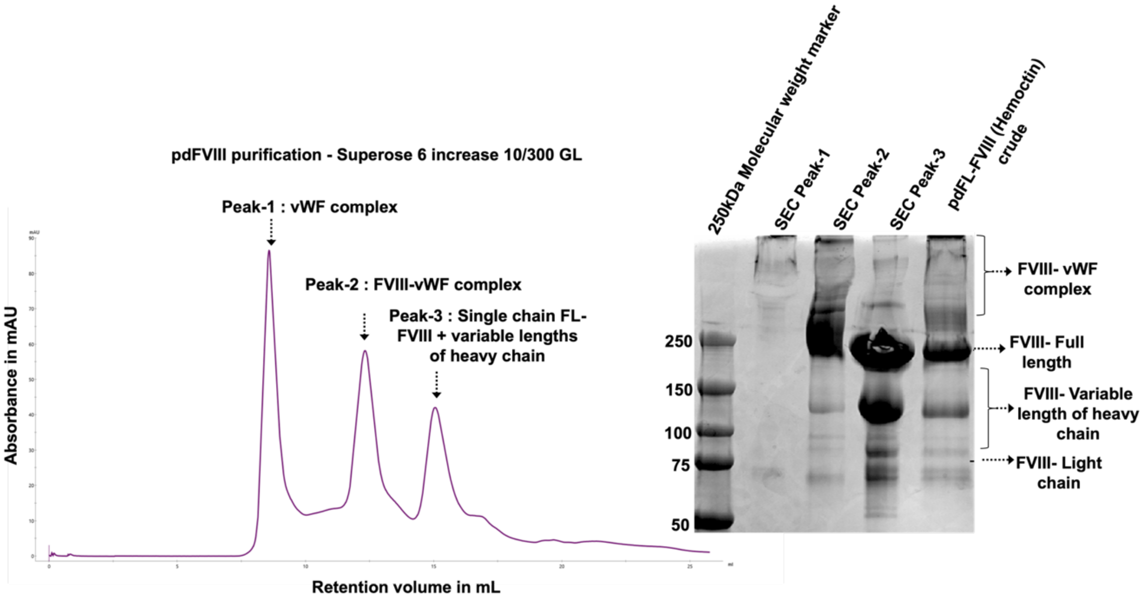


Figure 24B. SEC purification of pdFVIII. SEC chromatogram of pdFVIII performed using a Superose 6 Increase 10/300 GL column (Left). The chromatogram displays three distinct peaks: Peak 1 corresponds to the vWF complex, Peak 2 represents the FVIII-vWF complex, and Peak 3 corresponds to single-chain full-length FVIII with variable lengths of the heavy chain. SDS-PAGE analysis of SEC fractions under reducing conditions. Lane 1 shows the molecular weight marker, followed by SEC fractions (Peaks 1, 2, and 3), and a purified pdFVIII crude. The gel demonstrates the presence of vWF in Peak 1, FVIII-vWF complex in Peak 2, and full-length FVIII along with variable heavy chain fragments in Peak 3 (Right). Molecular weights are indicated on the left side of the gel.

SEC purification of pdFL-FVIII: The purification of pdFL-FVIII was performed using a Superose™ 6 Increase 10/300 GL (Cytiva) column based on its molecular size distribution and complexation with vWF. The chromatogram displayed three distinct peaks, corresponding to different molecular species in the sample (Figure 24B).

- **Peak-1:** This major peak represented the high-MW (HMW) vWF multimer (MM) complexes. Its early retention volume ~ 9.5 mL was consistent and aligned with the expected MW of high-MW multimers (HMW-MM) - above 300-500 kDa observed in pdFL-FVIII. The sharp and symmetrical profile of this peak indicates a homogeneous population of HMW-MM. The higher absorbance intensity of above 85 mAU for Peak-1 suggested that it constituted the major component of the pdFL-FVIII product (Figure 24B).
- **Peak-2:** This intermediate peak corresponded to vWF intermediate-lower-MW multimers (I-LMW-MM) and FVIII-vWF complexes which were smaller in size compared to Peak-1. Its intermediate retention volume of ~ 12 mL was consistent and aligned with the expected MW of I-LMW-MM and FVIII-vWF complex of 250 - 300 kDa observed in pdFL-FVIII. The presence of this peak highlighted the uniform distribution of FVIII molecules in various states of vWF MM and FVIII-vWF

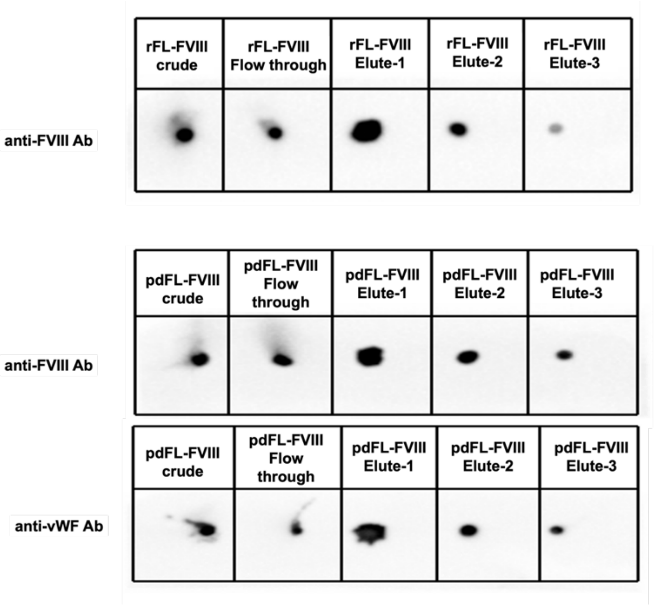
complexes. The higher absorbance intensity of above 50 mAU for Peak-2 suggested that the FVIII-vWF complex constituted the major component of the pdFL-FVIII product (Figure 24B).

- **Peak-3:** This peak represented the elution of sc-FL-FVIII, along with FVIII variants with variable lengths of HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$. The broader profile of Peak-3 and its retention volume of later than ~15 mL reflected heterogeneity in the population. The later elution volume of this peak was consistent and aligned with the LMW of the FVIII protein. The higher absorbance intensity of above 40 mAU for Peak-3 suggested that FVIII protein constituted the major component of the pdFL-FVIII product. The elution profiles demonstrated the efficiency of SEC in separating vWF HMW-MM complexes from FVIII variants. Meanwhile, the relative intensity and distribution of Peak-3 implied partial dissociation of FVIII from vWF (Figure 24B).

The separation of pdFL-FVIII proteins achieved ed by Superose™ 6 Increase 10/300 GL (Cytiva) column highlighted the efficiency of SEC in isolating intact vWF-MM and FL-FVIII forms from its variable MW/complex forms. SDS-PAGE analysis and coomassie staining of pdFL-FVIII revealed SEC Peak-1 of MW above 300-500 kDa was not resolved in the normal gradient gel of 4-12 % confirming its identity as the HMW-MM. Additionally, Peak-2 showed bands representing the I-LMW-MM and FVIII-vWF complex of 250 - 300 kDa. In contrast, SEC Peak-3 showed bands representing variable lengths of HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$, of ~ 150-200 kDa and ~ 75 kDa respectively. The unpurified crude pdFL-FVIII displayed a mixture of multiple bands of vWF and FVIII proteins (Figure 24B).

3.2.1.B Purification of rFL-FVIII, and pdFL-FVIII by co-IP and IEX

- rFL-FVIII and pdFL-FVIII were successfully purified by co-IP using anti-FVIII and anti-vWF Ab as detailed in Section 2.3.2.B. The dot blots analyzed using anti-FVIII Ab for rFL-FVIII purification displayed strong intensity dots in the first elutions and weaker intensities with the other 2 elutions. The flow-through also displayed a medium signal, indicating inefficient binding and loss of sample during the purification (Figure 25A). For, pdFL-FVIII purification, a similar observation was made for anti-FVIII and anti-vWF Ab blots with the strongest signal in the earlier elutions followed by weaker signals for the later elutions. Additionally, dot blots analyzed using anti-vWF Ab also displayed a similar pattern of stronger to weaker intensities, suggesting a strong association of FVIII-vWF complex in the pdFL-FVIII products (Figure 25B).



Samples		Description	S	SC	F
C1	P12263*	*Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	525	2,3	186
C2	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	390	2,8	160
C3	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	151	1,2	79
C4	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	413	3,5	162
C5	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	1.293	10,7	417
C6	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	510	4,5	194
C7	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	375	3,9	226
F1	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	382	2,2	106
F2	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	169	2,8	101
F4	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	39	2,5	57
F5	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	644	6,5	257
F6	P12263	Coagulation factor VIII OS=Sus scrofa OX=9823 GN=F8 PE=1 SV=2	37	2,3	42

Figure 25A. Dot blot analysis of rFL-FVIII and pdFL-FVIII purifications from co-IP. The anti-FVIII antibody panel detects rFL-FVIII in crude, flow-through, and elution fractions (Elute-1, Elute-2, and Elute-3), confirming its presence and enrichment of FVIII during purification. Similarly, pdFL-FVIII fractions (crude, flow-through, and elution fractions) were detected with the anti-FVIII antibody. Detection with the anti-vWF antibody reveals co-purification of vWF with pdFL-FVIII in crude and flow-through fractions, but minimal signal in elution fractions, indicating effective separation during purification. Mass spectrometry analysis confirmed the purity of the FVIII proteins (Right).

- For the IEX experiments, as mentioned in Section 2.3.2.B a gradient of $CaCl_2$ was employed to purify and yield high amounts of the FVIII protein. Dot blots corresponding to the purified protein are shown in Figure 25B. The FVIII-vWF complex was separated into two distinct peaks, with FVIII eluting at approximately 250 mM $CaCl_2$ and vWF at 300 mM, confirmed by SDS-PAGE analysis of the fractions. This indicated that FVIII and vWF bind in a Ca^{2+}-dependent manner, with the presence of calcium ions being critical for their interaction and stability. The eluted fractions were also subjected to functional assays, confirming that FVIII retained its activity post-separation, as evidenced by its ability to support coagulation in an in vitro assay. These results highlight the essential role of Ca^{2+} in maintaining FVIII-vWF complex stability, providing insights into their functional relationship in blood coagulation.

All the proteins purified through SEC, co-IP, and IEX were pooled, and concentrated to obtain high amounts of FVIII protein for further biophysical studies.

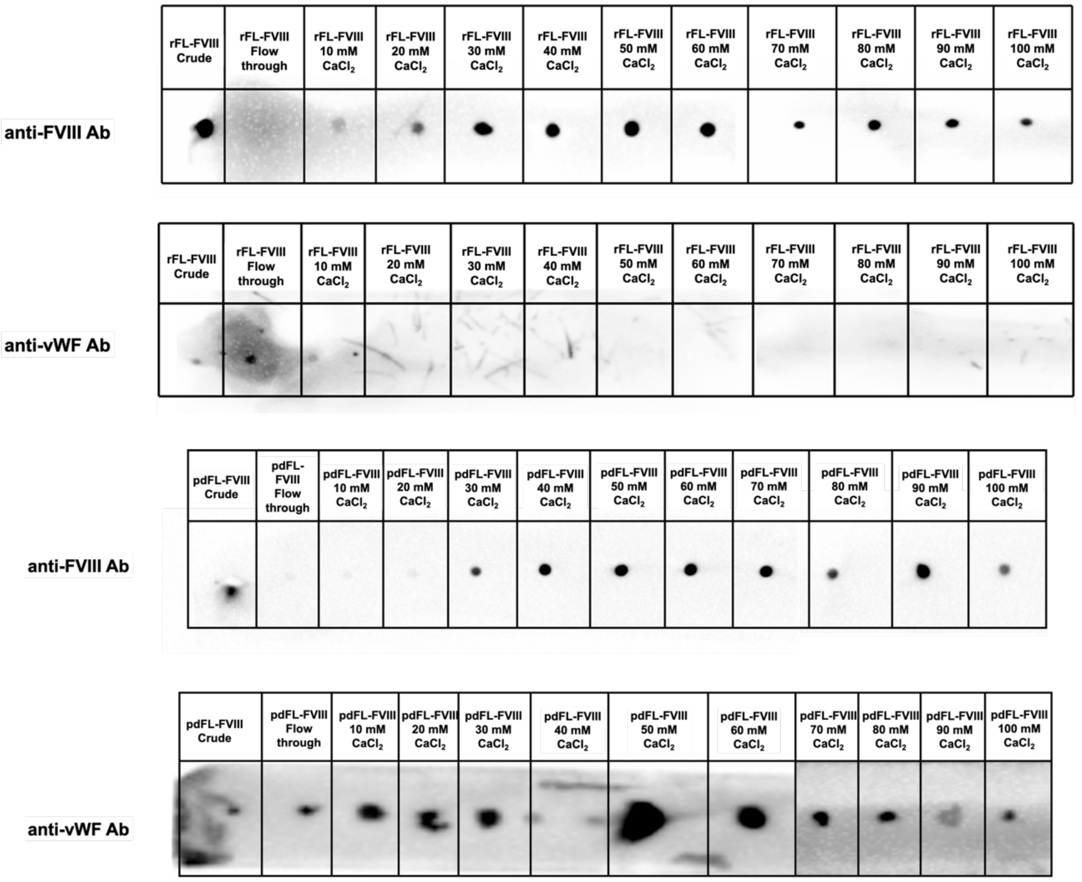


Figure 25B. Dot blot analysis of rFL-FVIII and pdFL-FVIII purifications from IEX. Dot blot analysis of rFL-FVIII (Top) pdFL-FVIII purification (Bottom) using anti-FVIII and anti-vWF antibodies. Anti-FVIII antibody detection shows FVIII protein presence in crude, flow-through, and elution fractions with increasing $CaCl_2$ concentrations. While anti-vWF antibody detection confirms vWF co-purification in crude and flow-through fractions, with reduced signals in elution fractions, indicating effective FVIII-vWF separation in pdFL-FVIII and minimal detection in the flow through in the rFL-FVIII proteins.

3.2.2 FVIII protein characterization studies - FVIII activity assay and vWF multimer analysis

	Peak-1	Peak-2	Peak-3
rFL-FVIII	FVIII Ac: 57.1 %	FVIII Ac: 10.8 %	-
pdFL-FVIII	vWF Ag: 458.2 % FVIII Ac: <<<	vWF Ag: 52.4 % FVIII Ac: 27.3 %	vWF Ag: 7.6 % FVIII Ac: 61.8 %

Table 10. FVIII activity assay of the purified rFL-FVIII and pdFL-FVIII proteins.

The analysis of rFL-FVIII and pdFL-FVIII revealed distinct profiles of FVIII Ac and vWF Ag across chromatographic peaks. For rFL-FVIII, Peak-1 displayed the highest activity (57.1

%), with significantly reduced activity in Peak-2 (10.8 %) and no significant measurable activity in Peak-3, suggesting the primary functional fraction resides in Peak-1. In contrast, pdFL-FVIII showed a markedly different distribution, with Peak-1 enriched in vWF antigen (458.2 %) but minimal FVIII activity, while Peak-2 exhibited moderate levels of both vWF Ag (52.4 %) and FVIII Ac (27.3 %). Notably, Peak-3 demonstrated the highest FVIII activity (61.8 %) for pdFL-FVIII, despite having minimal vWF Ag (7.6 %), indicating that functional FVIII in plasma-derived preparations may be more concentrated in later eluting fractions (Table 10).

3.2.3 Protein Biochemistry studies - vWF multimer analysis

The multimer analysis focused on evaluating the SEC-purified peaks of pdFL-FVIII associated with the vWF complex. The SEC profile and SDS-PAGE analysis of pdFL-FVIII resolved into three distinct peaks:

- Peak-1 represented the high-molecular-weight vWF complex,
- Peak-2 corresponded to the FL-FVIII-vWF complex, and
- Peak-3 contained a single chain and variable lengths of heavy chain and light chain of FVIII protein.

vWF-MM analysis provided insights to characterize the multimer composition and to understand what type of vWF complex multimers usually bind to FVIII protein, conducted using normal pool plasma (NPP) as a reference for high to low MW vWF multimers. Comparison with the NPP multimer profile highlighted the distinct distribution of multimers across the peaks, with Peak-1 resembling HMW multimers, Peak-2 reflecting LMW multimeric fractions bound to FVIII, and Peak-3 displaying no vWF multimers.

Peak-1 exhibited a significant proportion of high MW multimers (HMW-MM), with percentages of 16.3 %, 14.9 %, and 12.9 % corresponding to octadecamers to decamers, reflecting the dissociation of larger vWF multimers from FVIII (Figure 26).

Peak-2 demonstrated enrichment in middle-lower MW multimers (M-LMW-MM), with 22.2 %, 23.1 %, and 15.0 % associated with FVIII binding to middle-low MW vWF multimers (Figure 26).

Meanwhile, **Peak-3** was composed almost exclusively of purified FVIII with 100 % of minimal LMW-MM or no multimers, confirming the effective separation of FVIII from vWF. Multimer gels ran simultaneously and also provided information on the distribution and size of vWF multimers in peak-1, peak-2, and peak-3 samples. A low-resolution agarose gel loaded with the purified peaks and vWF multimers was separated based on their size. corresponding to peak-1 displayed high-molecular-weight (HMW) multimers, peak-2 displayed intermediate-sized multimers, and peak-3 displayed low-molecular-weight (LMW) fragments (Figure 26).

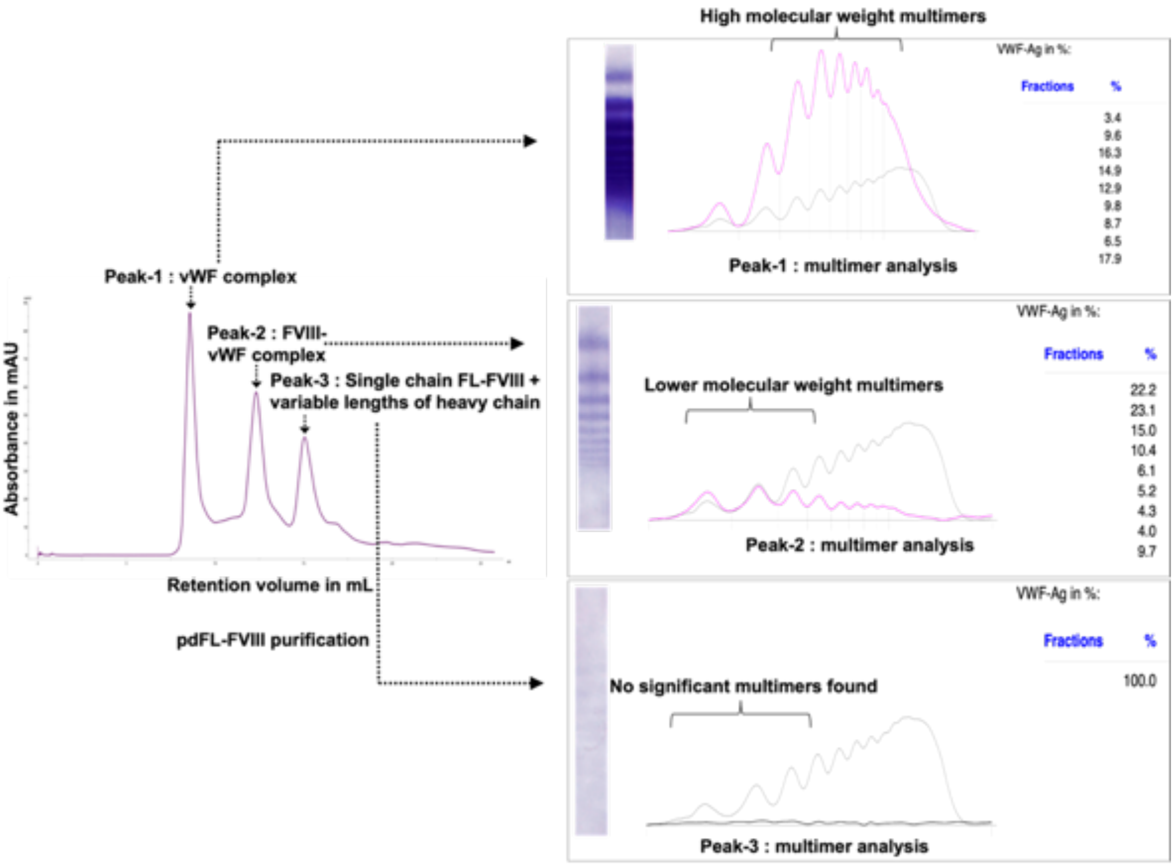
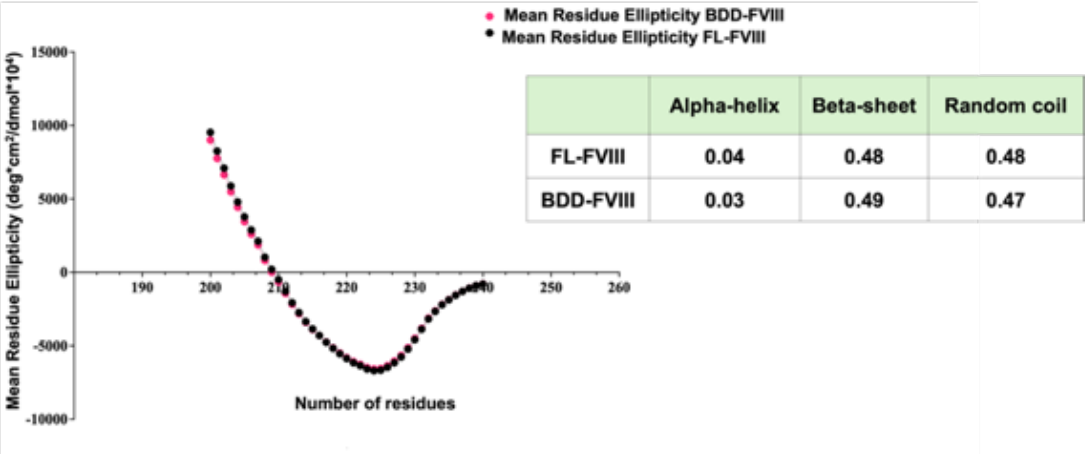


Figure 26. Multimer analysis of pdFL-FVIII. The variable peaks obtained during the purification of pdFL-FVIII were subjected to multimer analysis to confirm the peak-2 containing the FVIII-vWF complex. The graph of peak-1 corresponding to vWF complex (pink curve) analysis shows the presence of high molecular weight multimers and their respective percentage is also represented in the same, while the graph of peak-2 corresponding to FVIII-vWF complex (pink curve) analysis shows the presence of higher percentage of low molecular weight multimers (dimers and trimers) and there are no significant multimers found in peak-3 corresponding to just FVIII protein. All the peaks were compared to analysis performed on normal pool plasma (black curve) The multimer analysis gel images are represented on the top left of each section, showing the presence of high-low molecular weight multimers in a respective peak analysis.

3.2.4 Protein Biochemistry studies - CD studies

Purified proteins corresponding to 0.71 µg/µL (0.26 µM) and 0.27 µg/µL (0.16 µM) for FL-FVIII and BDD-FVIII were used, respectively for the CD studies. The UV-CD spectra were measured in 1X PBS at pH 7.4 and room temperature (21°C) and showed almost identical spectra for the rFL-FVIII and rBDD-FVIII in the 220-260 range as mentioned in Section 2.3.10. rFL-FVIII had a slightly broader minimum due to the presence of partial or full B-domain in the heavy chain containing variable mass FL-FVIII heterodimers. The calculated secondary structure distribution showed a predominant beta-strand content of ~48 % and a low alpha-helical content of ~4 %. These similarities in the secondary structure distribution between the rFL-FVIII and rBDD-FVIII show that the B-domain because of its disorderness has little to no effect on the overall FVIII secondary structure (Figure 27). The secondary structure content was observed for simulated and refined computational *Gly-BDD-FVIII* and *Gly/FC-FL-FVIII* models. For, *Gly-BDD-FVIII* - 4 % Helix, 39.41 % Sheet, 18 % Turn, and 42 % Coils were observed. While *Gly/FC-FL-FVIII* - 6.37 % Helix, 28.10 % Sheet, 14 % Turn, and 60 % Coil was observed. Slight differences in the secondary structure content between *Gly-BDD-FVIII* and *Gly/FC-FL-FVIII* were found. Additionally secondary structure content for available BDD-FVIII crystal structure PDB-ID: 3CDZ was also calculated showing 1.67 % Helix, 39.41 % Sheet, and 58.92 % other secondary structure contents. Even though slight differences between all three structures were observed, it is notable to consider that the *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII* are simulation equilibrated and refined computational models and 3CDZ is a crystal static structure adopted from RCSB-PDB depicting that crucial secondary structures could be captured in a detailed manner in the latter (Figure 27).



	Alpha-helix	Beta-sheet	Random coil
FL-FVIII	0.04	0.48	0.48
BDD-FVIII	0.03	0.49	0.47

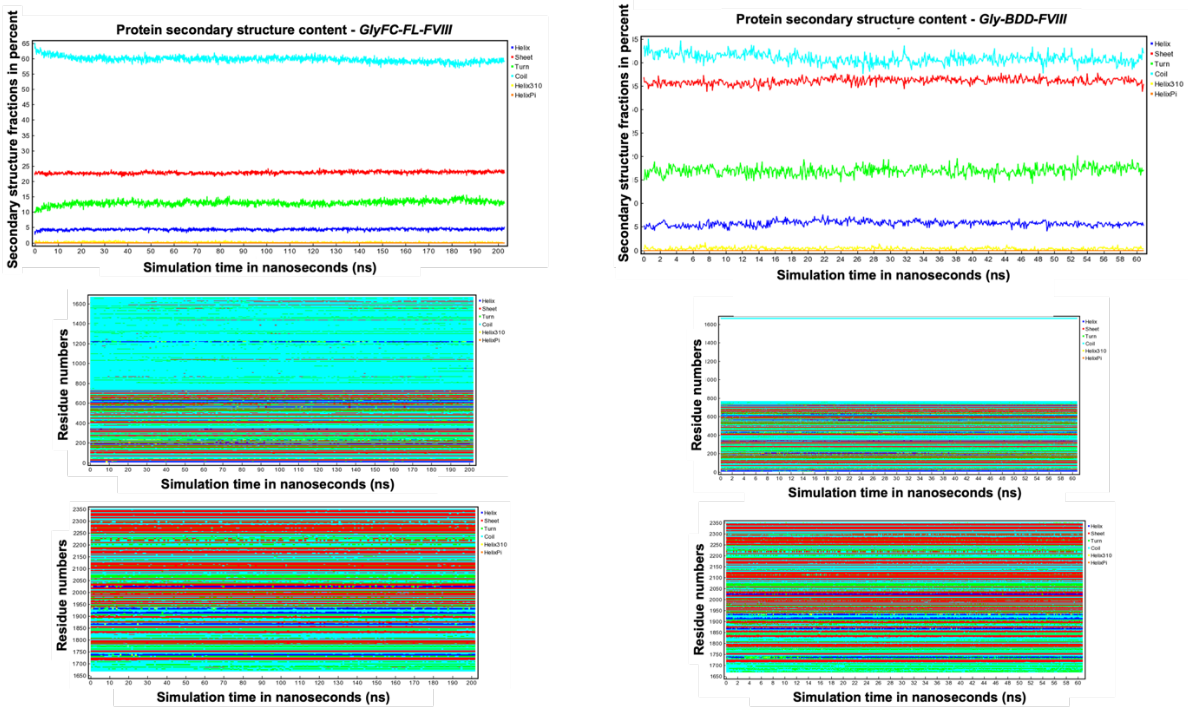


Figure 27. CD studies and comparison of secondary structures between rFL-FVIII and rBDD-FVIII. CD graph compares the difference in the secondary structure content observed as mean residue ellipticity (MRE) of FL-FVIII, (black) and BDD-FVIII, (pink). The table shows the values corresponding to the secondary structure analysis from Dichroweb [135] for both proteins, with minor differences in alpha-helix, beta-sheet, and random coil proportions. Secondary structure content between the simulation refined *Gly-BDD-FVIII* and *Gly/FC-FL-FVIII* computational models displayed variable contents of the helix (dark blue), turns (green), sheet (red), and coils (cyan blue) are displayed in the graphs.

3.2.5 Protein Biochemistry studies - Deglycosylation studies

Deglycosylation of rFL-FVIII protein using PNGase F was analyzed on SDS-PAGE and coomassie staining. A slight electrophoretic mobility shift of the single chain band at 250 kDa of rFL-FVIII was observed due to the removal of glycosylation. Similarly, the second trial of deglycosylation studies also displayed a shift in the variable lengths of HC $_{\text{(A1-A2-domains)}}$ between 150 kDa and 100 kDa and LC $_{\text{(A3-C1-C2-domains)}}$ at 75 kDa. The extent of deglycosylation was observed using mass spectrometry studies. Deglycosylated Asn residues underwent deamidation modification and this was followed in mass spectrometry studies where 9 glycosylation sites of N258, N776, N1020, N1074, N1301, N1431, N1829, N2137 were identified to be deamidated and deglycosylated (Figure 28). The combined analysis highlighted the detailed structural modifications resulting from the enzymatic removal of N-linked glycans, providing insight into the glycosylation profile and its functional impact on rFL-FVIII.

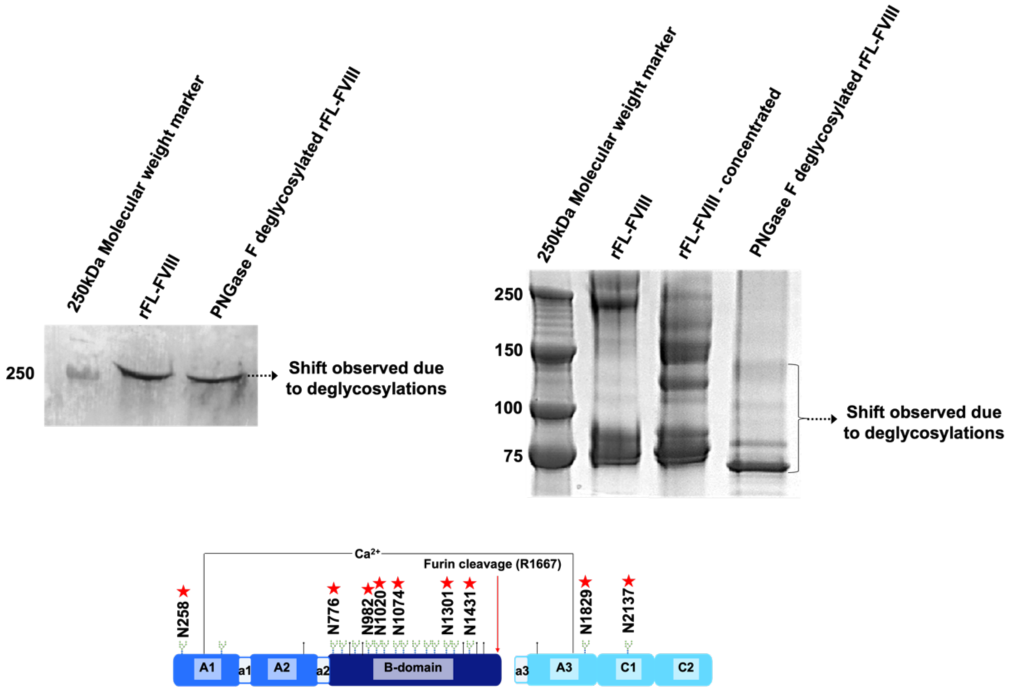


Figure 28. Deglycosylation and mass spectrometry studies of rFL-FVIII. Coomassie staining of deglycosylated rFL-FVIII displayed a slight electrophoretic mobility shift at 250 kDa and HC $_{\text{(A1-A2-domains)}}$ between 150 kDa and 100 kDa and LC $_{\text{(A3-C1-C2-domains)}}$ at 75 kDa. Domain structure with deamidated Asn residues after deglycosylation is also depicted below.

3.3 Results - Biophysical studies

3.3.1 Biophysical studies - Negative staining and cryo-electron microscopy studies

- **Negative staining:** Negative staining micrographs of purified pdFL-FVIII peaks revealed significant particle heterogeneity and conformational variability, indicating the dynamic nature of these complexes. Micrographs corresponding to peak-2 primarily displayed elongated, tail-like appendage structures that were well-dispersed and exhibited minimal aggregation. These appendages were measured to be approximately 20 - 30 nm in diameter, reflecting the dimensions of tail-like elements of individual vWF multimers. Additionally, scattered among these appendages were globular particles, likely the FL-FVIII-vWF complex/ individual FVIII. In contrast, micrographs of peak-3 revealed predominantly globular particles with diameters ranging between 10 - 12 nm. These dimensions align with the expected size of FL-FVIII in its more compact form, suggesting that this peak largely consists of FL-FVIII molecules without significant association with vWF multimers. The clear distinction in morphology and size between the particles in peaks 2 and -3 highlighted reduced uniformity and high heterogeneity, making them unsuitable for further 2D class averaging and 3D reconstruction (Figure 29).

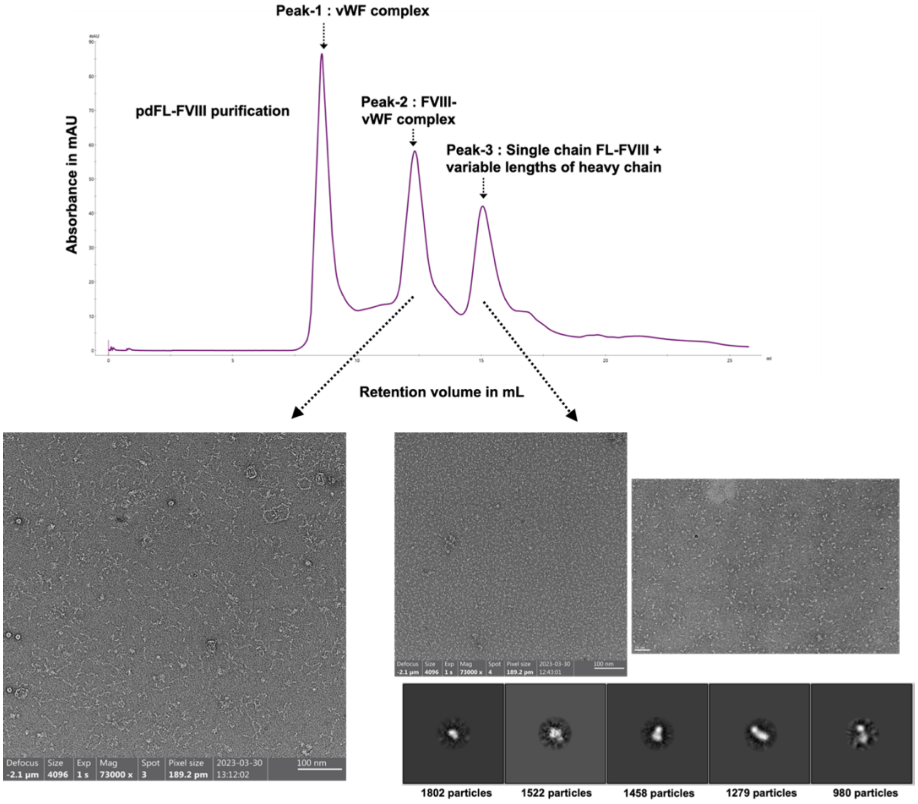


Figure 29. Negative staining studies of rFL-FVIII and pdFL-FVIII proteins. The SEC-purified rFL-FVIII and pdFL-FVIII displayed distinct peaks depicted in the top panel. Peak-1 corresponds to the vWF complex, Peak-2 represents the FVIII-vWF complex, and Peak-3 contains single-chain FL-FVIII with variable lengths of the heavy chain. The lower panel presents representative negative-staining images of the particles from each peak. Raw micrographs display particle distribution and integrity across the elution peaks, with corresponding 2D class averages shown at the bottom. Particle counts for each class average are indicated, with 1982 particles, 1602 particles, 1348 particles, 1278 particles, and 980 particles contributing to the analysis.

Further particle counts of individual cropped micrographs displayed particle densities as follows:

Micrograph 1 exhibited the highest density with 1,802 particles
Micrograph 2 with 1,522 particles
Micrograph 3 displayed 1,480 particles
Micrograph 4 showed the lowest density, containing 968 particles

The micrographs provided a detailed visualization of the sample composition. Micrograph 1, with its higher particle count, displays a densely populated field, while, Micrograph 4, with the lowest particle count, suggested a sparser distribution of particles. These counts reflect the diverse particle distribution and structural variability within the sample, suggesting significant heterogeneity across the observed fields (Figure 29).

- **Cryo-electron microscopy (cryo-EM)**

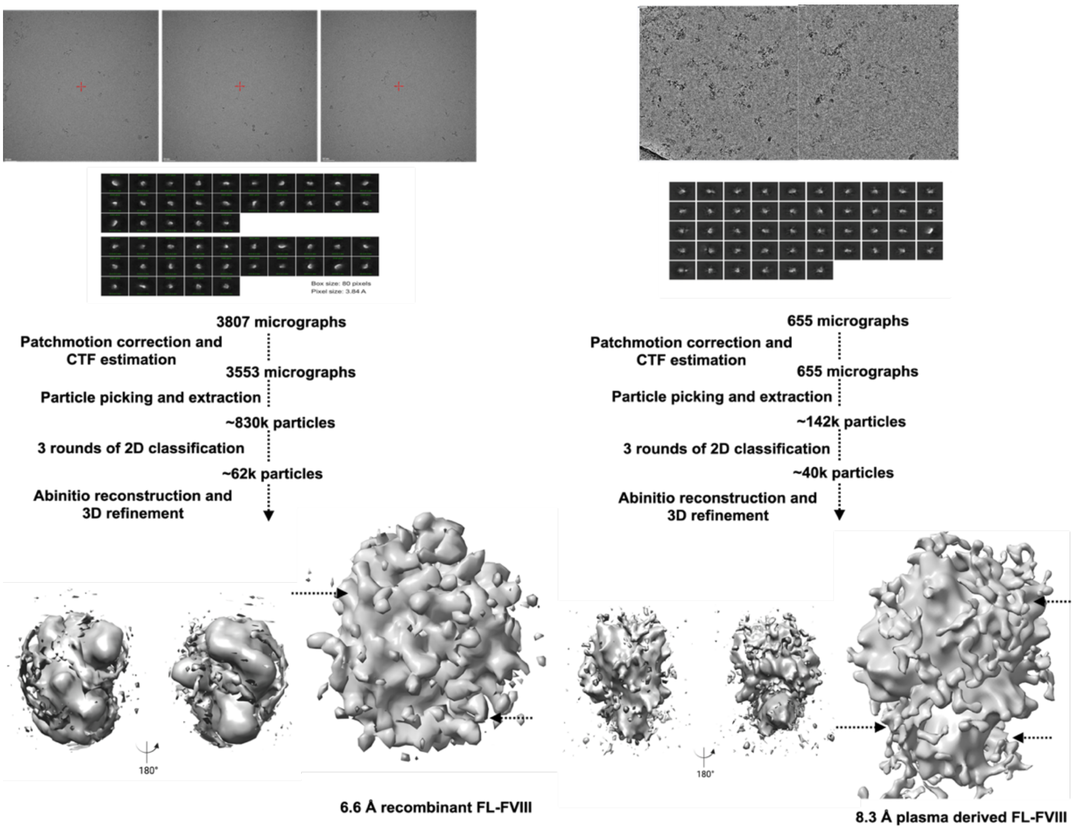
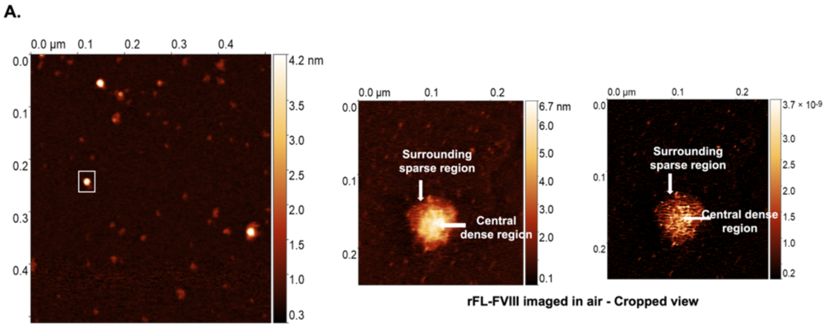


Figure 30. Low-resolution cryo-EM structure of rFL-FVIII and pdFL-FVIII protein. 2D classification profile (top left) and density maps of rFL-FVIII (below left) ≈ 7 Å rFL-FVIII structures. The B-domain was found to be surrounding the central core (marked in arrows). Similarly, pdFL-FVIII 2D classification profile (top right) and density maps of pdFL-FVIII (below right) ≈ 8 Å rFL-FVIII structures. The B-domain was found to be surrounding the central core (marked in arrows).

3807 movies for rFL-FVIII were collected with aberration-free image shift (AFIS), with a pixel size of 0.93 Å, a defocus range of - 0.9 to - 1.8 µm, and a total dose of 40 $e^-/Å^2$. Cryo-EM data processed via cryoSPARC v3.3.2 were primarily motion-corrected and a patch-based CTF was estimated. After motion correction, ~ 830,000 particles were picked which were then subjected to 3 rounds of 2D classification. The remaining ~ 62,000 particles were used for ab initio reconstruction and 3D refinement. Similarly, 655 movies for pdFL-FVIII were collected with a pixel size of 0.727 Å and a total dose of 50 $e^-/Å^2$. After motion correction, 370,000 particles were picked using the blob-picking algorithm of cryoSPARC. Three rounds of 2D classification yielded ~ 142,000 particles for analysis, which were picked and subjected to 2D classification, for Ab-initio reconstruction with four classes. One class with 40,000 particles was subjected to 3D Homogenous Refinement for final 3D model reconstruction. Final 3D refinements and reconstructions yielded approximately ~6.6 Å resolution density map for rFL-FVIII and ~8.3 Å resolution density map for pdFL-FVIII. Both density maps were primarily globular, with sparse densities observed surrounding the central globular regions (Figure 30). Due to the low resolution of these maps, direct model building was not feasible. As a result, rigid body fitting, as further detailed in Section 3.4.1, was employed to position and validate the ordered domains of the structure within the density maps.

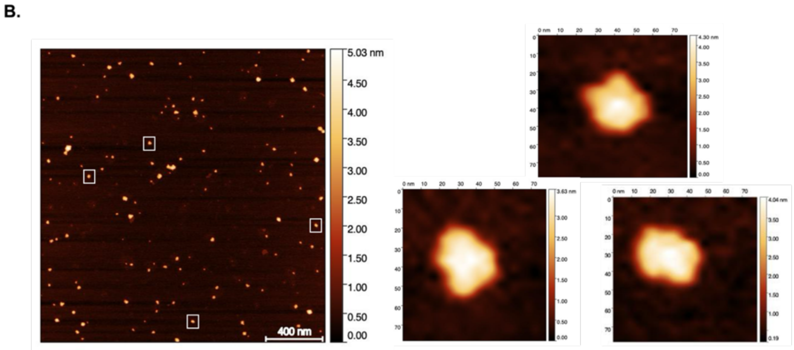
3.3.2 Biophysical Studies - Atomic Force Microscopy

- **Air AFM:** Air AFM analysis of rFL-FVIII and rBDD-FVIII proteins revealed more globular structures. The rFL-FVIII, globular regions displayed a height of ~ 6-7 nm and a width of ~ 30-40 nm. Certain rFL-FVIII topographs revealed the presence of high dense region in the center surrounding a light-scattered region extended in different directions in each isolated surface. The rBDD-FVIII displayed a height of ~ 3-4 nm and a width of ~ 20-30 nm. Even though it also appeared to be a globular structure, certain distinct domains were visible in the cropped topographs (Figure 31).
- **Liquid AFM:** Liquid AFM analysis also revealed similar results for rFL-FVIII, globular regions displayed a height of ~ 6-7 nm and a width of ~ 30-40 nm. Whereas globular regions of rBDD-FVIII displayed a height of ~ 4-5 nm and a width of ~ 20-30 nm. Considering the difference between the two is the presence and the absence of the B-domain, it indicated that the additional height and width could have been contributed by the B-domain. Similarly, liquid AFM analysis of pdFL-FVIII revealed a height of ~ 6-7 nm for the globular regions, and a height of ~ 3-4 nm for tails-like regions The total width of pdFL-FVIII was found to be ~ 60-80 nm (Figure 31). The analysis was repeated three times, yielding consistent results with stable height and width measurements across rFL-FVIII, rBDD-FVIII, and the pdFL-FVIII-vWF complex.



rFL-FVIII imaged in air - Cropped view

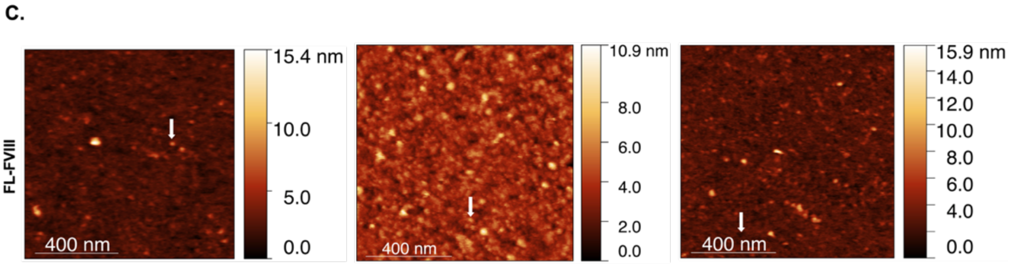
rFL-FVIII imaged in air - Global view



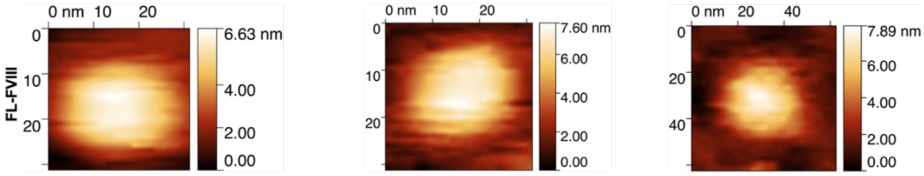
rBDD-FVIII imaged in air - Global view

rBDD-FVIII imaged in air - Cropped view

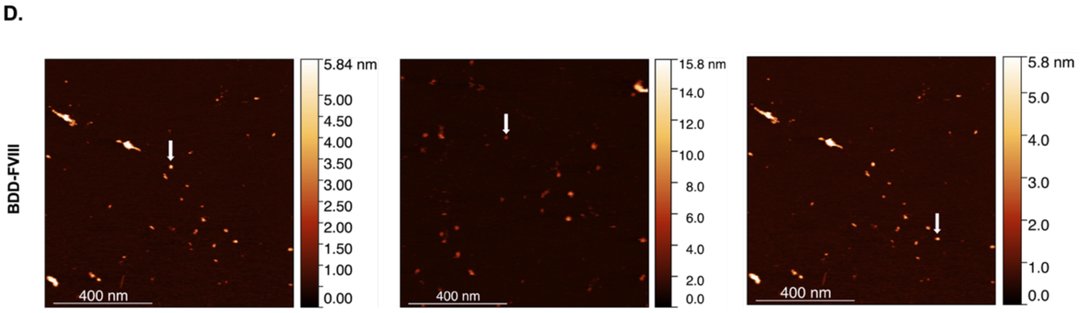
Figure 31. AFM studies of FL-FVIII, BDD-FVIII, and pdFL-FVIII-vWF complex. A. The global view (top left) displays an in-air AFM topograph of rFL-FVIII revealing globular structures, while the cropped view (top right) shows LPW-processed images of FL-FVIII, highlighting a central dense region surrounded by a sparsely scattered area. **B.** The global view (bottom left) displays an in-air AFM topograph of rBDD-FVIII revealing globular structures, while the cropped view (bottom right) shows LPW-processed images of rBDD-FVIII, highlighting only the presence of a central dense region.



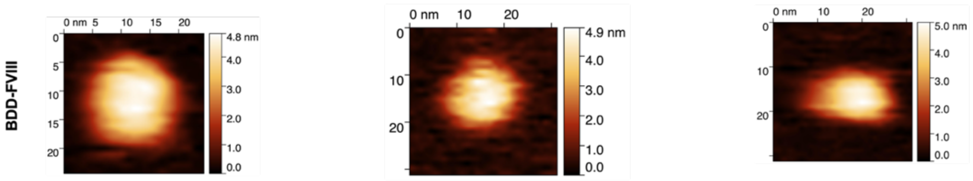
rFL-FVIII imaged in liquid - Global view



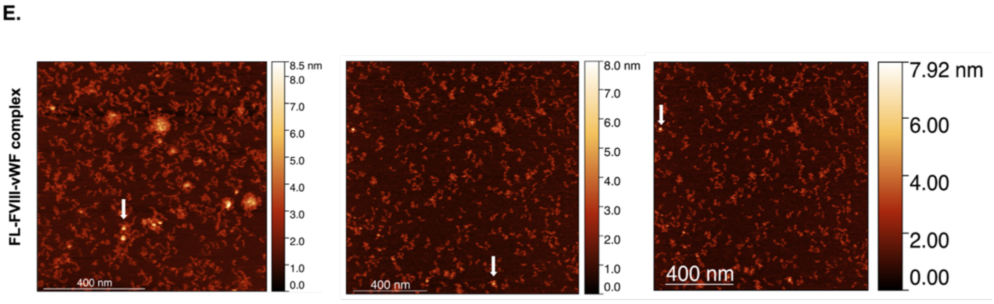
rFL-FVIII imaged in liquid - Cropped view



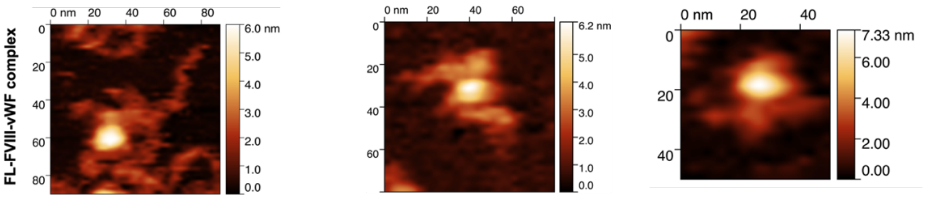
rBDD-FVIII imaged in liquid - Global view



rBDD-FVIII imaged in liquid - Cropped view



FL-FVIII-vWF complex imaged in liquid - Global view

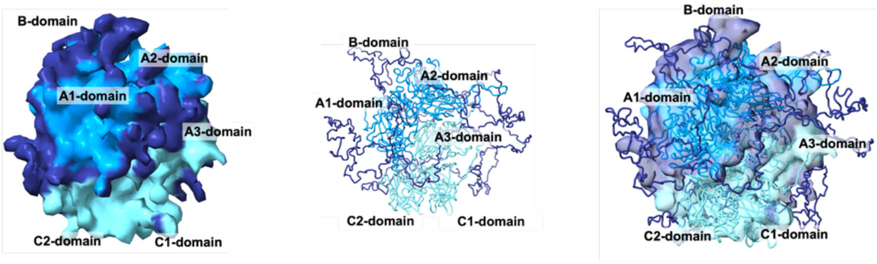


FL-FVIII-vWF complex imaged in liquid - Cropped view

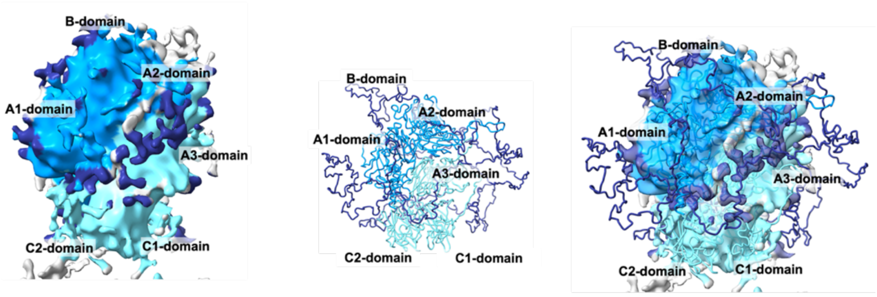
Figure 31. AFM studies of FL-FVIII, BDD-FVIII, and pdFL-FVIII-vWF complex. C. The global view of liquid AFM topographs of FL-FVIII (top panel) and **D.** BDD-FVIII (middle panel) show globular conformations and **E.** pdFL-FVIII-vWF complex (bottom panel) showing globular FVIII connected to vWF tails (towards left). White arrows in the global view represent the individual FL-FVIII, BDD-FVIII, and pdFL-FVIII-vWF individual molecules that were cropped. The cropped topographs on the right of rFL-FVIII and rBDD-FVIII showed a height of ~6-7 nm and ~4-5 nm, and a width of ~ 30-40 nm and ~ 20-30 nm respectively. The pdFL-FVIII-vWF complex showed a height of ~6-7 nm for the globular part, ~3-4 nm for tails, and a width of ~ 80-90 nm.

3.4 Results - Integrating biophysical studies data and computational models

3.4.1 Integrating biophysical studies data and computational models - Fitting computational models to low-resolution cryo-EM density maps



rFL-FVIII rigid body fitting



pdFL-FVIII rigid body fitting

Figure 32. Integrating computational models to low-resolution cryo-EM density maps. rFL-FVIII cryo-EM density map (≈ 6.6 Å) fitted to the simulation equilibrated *Gly/FC-FL-FVIII* model using PowerFit Server and ChimeraX (top panel). The specific domains and their relative spatial positions are marked showing the presence of the B-domain surrounding the core of FVIII protein. Similarly, the pdFL-FVIII cryo-EM density map (≈ 8 Å) fitted to the simulation equilibrated *Gly/FC-FL-FVIII* model using PowerFit Server and ChimeraX. The specific domains and their relative spatial positions are marked showing the presence of the B-domain surrounding the core of the FVIII protein (bottom panel).

Rigid body fitting carried out using the PowerFit server and the Fit-in-Map function of ChimeraX further refined the model alignment of *Gly/FC-FL-FVIII* model with rFL-FVIII and pdFL-FVIII cryo-EM density maps. Rigid body fitting of the cryo-EM map accommodated the ordered domains of FVIII - HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$ however, no distinct density was observed corresponding to the B-domain. Nonetheless, certain regions of the B-domain in the *Gly/FC-FL-FVIII* model overlapped with the sparsely distributed density regions surrounding it (Figure 32). The RMSD values were maintained below 2 Å for ordered domains for both rFL-FVIII and pdFL-FVIII cryo-EM density maps, which is generally considered optimal for a good fit, indicating that the model closely matched the experimental density in these ordered regions. Whereas the RMSD was way above for the B-domain fits. The Cross-Correlation Coefficient (CCC) quantified the correlation between the model and the maps, with values above 0.5 for the ordered regions and 0 for the B-domain fits. The real-space correlation score, which assesses how well the model follows the electron density in real space, also exhibited high values of 0.7, confirming a good model fitting for the ordered regions, whereas the B-domain revealed a value of 0.2 which indicated electron density in real space for B-domain were not distinct and geometry/positioning of the B-domain deviated significantly due to its flexibility. Additionally, visual inspection of the cryo-EM density map confirmed that the ordered regions of the model aligned with the high-density regions for HC $_{\text{(A1-A2-domains)}}$ and LC $_{\text{(A3-C1-C2-domains)}}$, further corroborating the fitting quality (Figure 32).

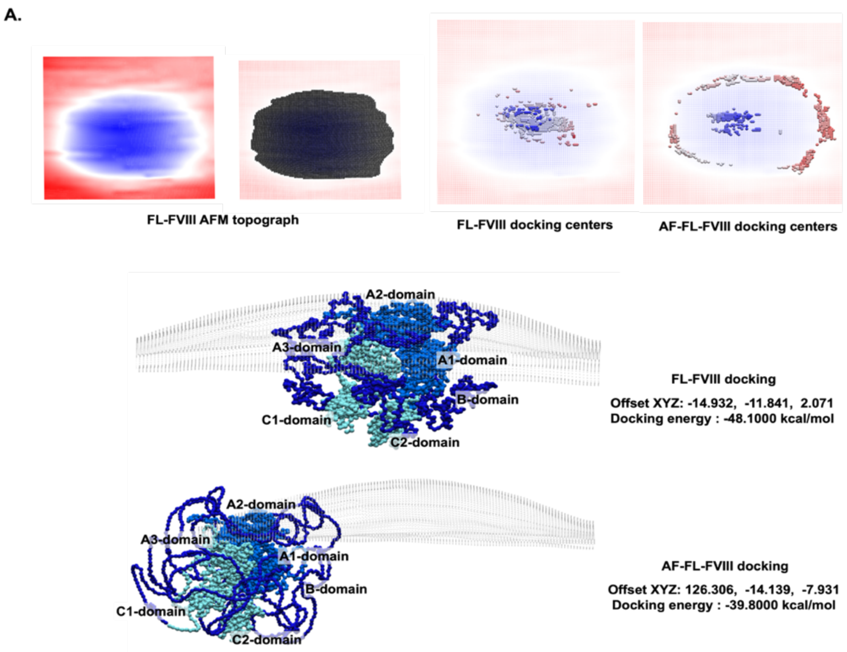
3.4.2 Integrating computational models to AFM topographs through DockAFM studies

The original height of all the FL-FVIII, FL-FVIII-vWF, and BDD-FVIII AFM topographs were flattened and denoised using Gwyddion detailed in Section 2.4.2. The de-noised images were cropped and adjusted using Gwyddion, and an exhaustive search for all possible docking orientations was carried out using the DockAFM/AFM Assembly pipeline. DOT2 enabled the translation of denoised and cropped 3D volumetric data for FL-FVIII, FL-FVIII-vWF, and BDD-FVIII models from their respective AFM topographs. The DOT2 software successfully performed a comprehensive six-dimensional search, involving translations along the X, Y, and Z axes and three Euler rotational degrees of freedom of the individual FL-FVIII, BDD-FVIII, and FL-FVIII-vWF AFM topographs. DOT2 also translated and rotated the 3D atomic coordinates of each *Gly/FC-FL-FVIII*, 3CDZ, and *Gly/FC-FL-FVIII-vWF-D'-D3* complex model. Further observations from comparative DockAFM studies revealed the quality of both AFM topographs obtained and refined models generated (Figure 33).

- **FL-FVIII assembly and docking:** Cropped FL-FVIII AFM topographs docked with refined and validated *Gly/FC-FL-FVIII* model and AF-FL-FVIII model showed better fitting with our refined *Gly/FC-FL-FVIII* model than compared to AF-FL-FVIII model. The *Gly/FC-FL-FVIII* model docked onto FL-FVIII topographs showed the lowest docking energy of -48.1000 kcal/mol with minimal XYZ displacement of -14.932, -11.841, and 2.071, respectively. Whereas the AF-FL-FVIII model showed a higher docking energy of -39.38000 kcal/mol with higher XYZ displacement of 126.306, -14.139, and -7.931, respectively. The top 20 docking orientations of the *Gly/FC-FL-FVIII* model were concentrated within the central docking layer region, whereas the AF-FL-FVIII model exhibited a dispersed distribution of docking centers and orientations (Figure 33A).
- **BDD-FVIII assembly and docking:** Cropped BDD-FVIII AFM topographs docked with refined and validated *Gly/FC-FL-FVIII* model and BDD-FVIII PDB - 3CDZ model and showed better fitting with *Gly-BDD-FVIII* model than compared to *Gly/FC-FL-FVIII* model. The 3CDZ model docked onto BDD-FVIII topographs showed the lowest docking energy of -48.4000 kcal/mol with minimal XYZ displacement of 2.999, 24.463, and 5.275 respectively. Whereas the *Gly/FC-FL-FVIII* model showed a higher docking energy of -32.3000 kcal/mol with higher XYZ

displacement of 16.463, -13.803, and -9.623 respectively. The top 20 docking orientations of the 3CDZ model were concentrated within the central docking layer region, whereas the *Gly/FC-FL-FVIII* model exhibited a dispersed distribution of docking centers and orientations. The Gly/FC-FL-FVIII model also appeared to be much larger to fit in the BDD-FVIII AFM topographs (Figure 33B).

- **FL-FVIII-vWF complex assembly and docking:** Similarly, cropped FL-FVIII-vWF AFM topographs docked with refined and validated *Gly/FC-FL-FVIII-vWF-D'-D3* complex model and *Gly/FC-FL-FVIII* model showed better fitting with *Gly/FC-FL-FVIII-vWF-D'-D3* complex model than compared to *Gly/FC-FL-FVIII* model. The *Gly/FC-FL-FVIII-vWF-D'-D3* model docked onto FL-FVIII-vWF topographs showed the lowest docking energy of 63.6000 kcal/mol with minimal XYZ displacement of -45.860, -47.865, -9.838 respectively. Whereas the *Gly/FC-FL-FVIII* model showed higher docking energy of -50.1000 kcal/mol with higher XYZ displacement of -53.079, -142.775, and -1.492 respectively. The top 20 docking orientations of the *Gly/FC-FL-FVIII-vWF-D'-D3* complex model were concentrated within the central docking layer region, whereas the *Gly/FC-FL-FVIII* model exhibited a dispersed distribution of docking centers and orientations. The *Gly/FC-FL-FVIII* model also appeared to be much smaller to fit in the FL-FVIII-vWF AFM topographs. *Gly/FC-FL-FVIII-vWF-D'-D3* also displayed D'-D3 regions extending towards the tail regions in FL-FVIII-vWF topographs (Figure 33C).



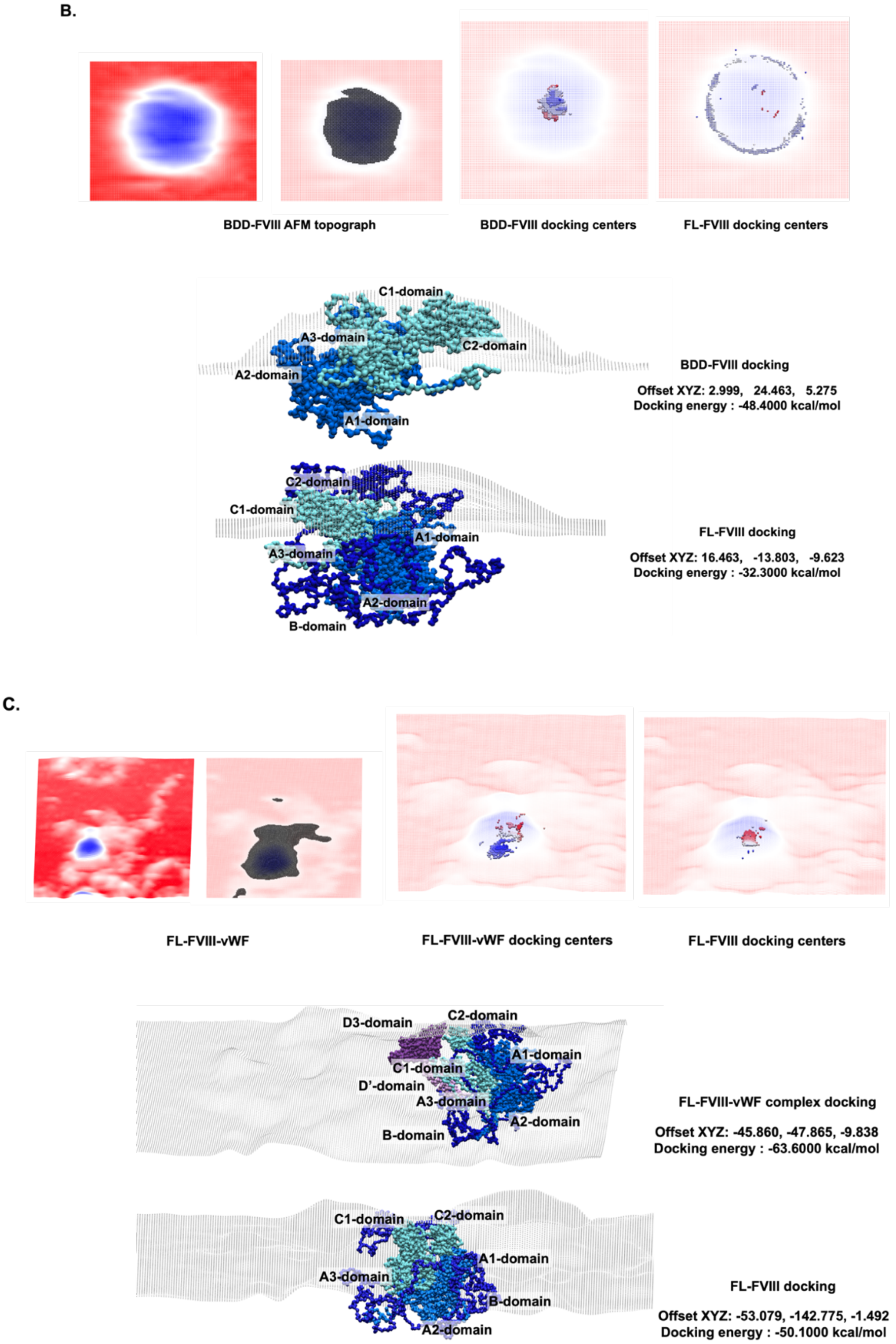


Figure 33. Integrating computational models to AFM topographs through DockAFM studies. DockAFM was performed to fit the structural models of *Gly/FC-FL-FVIII*, BDD-FVIII (PDB-ID: 3CDZ), and *Gly/FC-FL-FVIII-vWF-D'-D3* complex structural models to their respective AFM images (Figure 31). A comparative fitting was performed for individual images of **A.** FL-FVIII

images with *Gly/FC-FL-FVIII* and AF-FL-FVIII models, **B.** BDD-FVIII images with 3CDZ and *Gly/FC-FL-FVIII* models, and **C.** pdFL-FVIII-vWF images with *Gly/FC-FL-FVIII-vWF-D'-D3* complex and *Gly/FC-FL-FVIII* models. The images to the top left of each section display the cropped AFM topographs converted to 3D envelop with regions to be docked (marked in dark grey) The images in the center of each section show the comparative fitting of individual models to their respective images and the image on the top right corner show the best docking pose with minimum displacement and lowest docking energy for individual fitting. The offset XYZ and docking energy at the top right corner indicate the displacement and energy required for fitting individual models.

3.5 Results - Functional characterization and interaction studies

3.5.1 Functional characterization and interaction studies - Electrostatic studies

The Coulombic interaction values observed for *Gly/FC-FL-FVIII* reveal significant deviations from those calculated for the 3CDZ structure, primarily due to the influence of the B-domain and its inherent glycosylations. In the *Gly/FC-FL-FVIII* model, the HC $_{\text{(A1-A2-domains)}}$ exhibited a mean Coulombic value of +0.96 kcal/mol/e, contrasting with the 3CDZ HC $_{\text{(A1-A2-domains)}}$ value of -1.91 kcal/mol/e. Similarly, the LC $_{\text{(A3-C1-C2-domains)}}$ shifted from -0.26 kcal/mol/e in 3CDZ to +0.37 kcal/mol/e in *Gly/FC-FL-FVIII*. The positive shift in Coulombic potential for both chains in *Gly/FC-FL-FVIII* suggests that the B-domain adds regions of positive electrostatic potential (Figure 34). This shift in electrostatic potential highlighted the stabilizing effect of the glycosylated B-domain, which reduced the overall negativity of the molecule and created a more distributed influence on the local and long-range electrostatic environment. The positively charged regions, particularly residues such as R2333 and R2182 in the C-domains involved in PS membrane binding, retained their strong positive charge (Figure 34). The analysis further emphasized the orientation dynamics of *Gly/FC-FL-FVIII* when interacting with a negatively charged mica surface. During DockAFM studies as detailed in Section 3.4.2, we could observe several stable constant docking orientations of the *Gly/FC-FL-FVIII* model showing positively charged domains oriented towards negatively charged mica surface. This mimicked the binding of FVIII positively charged domains (C1-C2-domains) to PS membranes displaying the role of electrostatics in functionally important events (Figure 34).

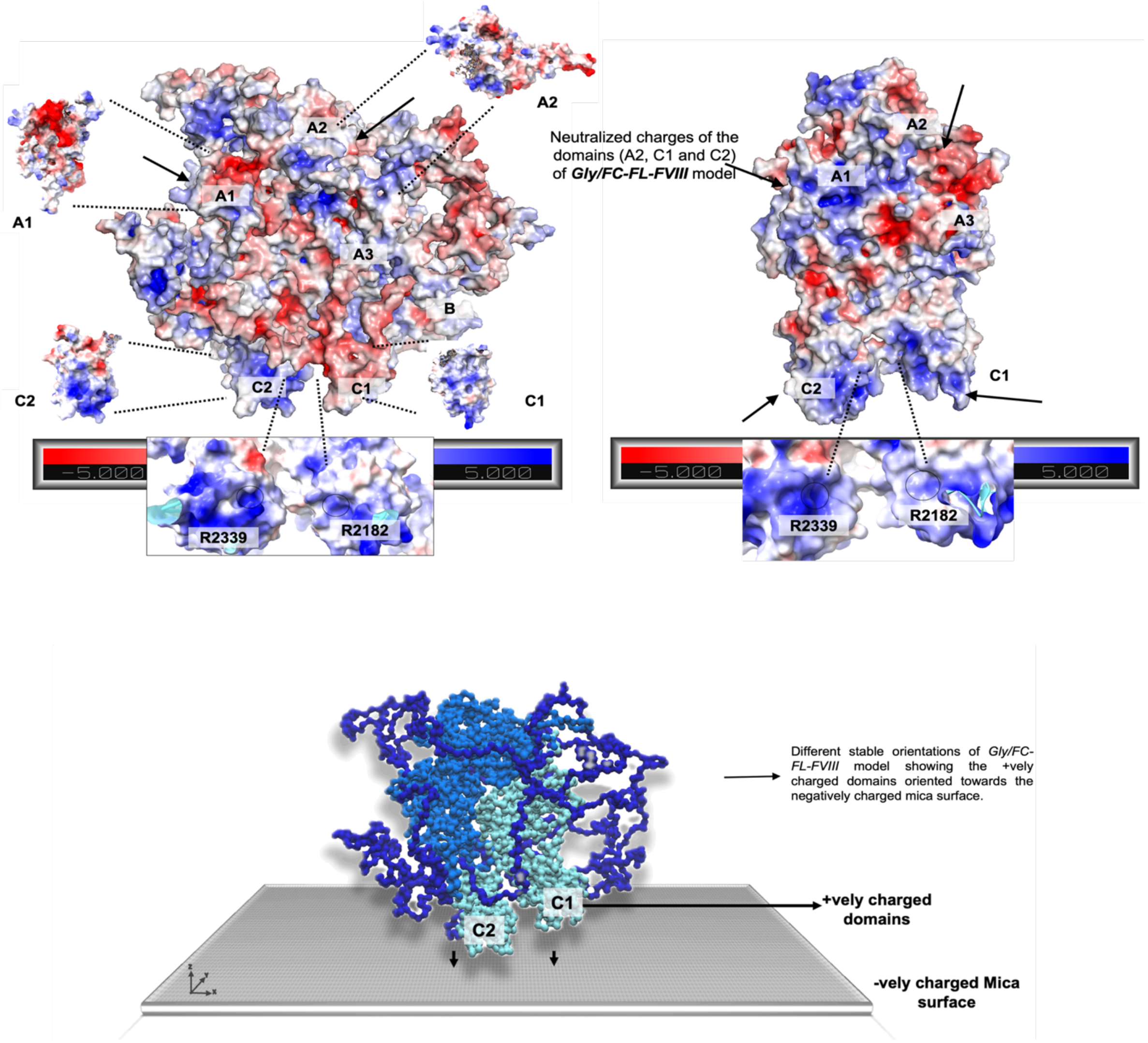
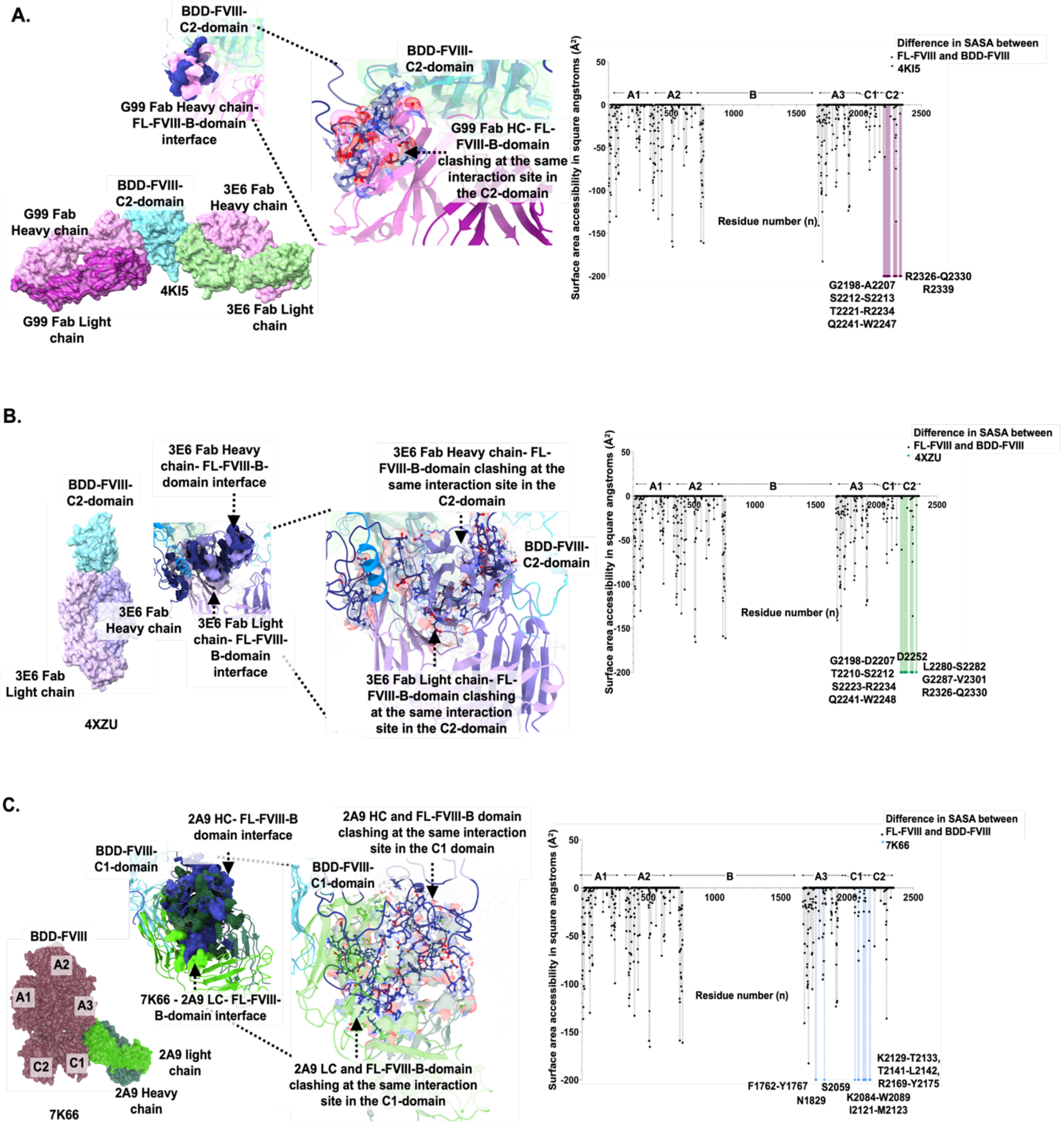


Figure 34. FVIII surface electrostatic studies. Electrostatics of *Gly/FC-FL-FVIII* and 3CDZ calculated using APBS show overall neutralized charges of the domains (A1-, A2-, C1- and C2-domains) in *Gly/FC-FL-FVIII* due to the presence of B-domain. The zoomed-in images show the positive patches retained in PS binding interfaces cantered on R2182 and R2339 (top panel). Docking of AFM topographs with the *Gly/FC-FL-FVIII* model displayed different stable orientations showing the +vely charged domains oriented towards the negatively charged mica surface. Both electrostatic studies reinforced how the presence of B-domain would affect binding to interacting partners, and inhibitors and influence premature activation and binding to PS surfaces (bottom panel).

3.5.2 Functional characterization and interaction studies - Surface area accessibility studies

3.5.2.A Surface area accessibility studies: FVIII Inhibitory antibodies



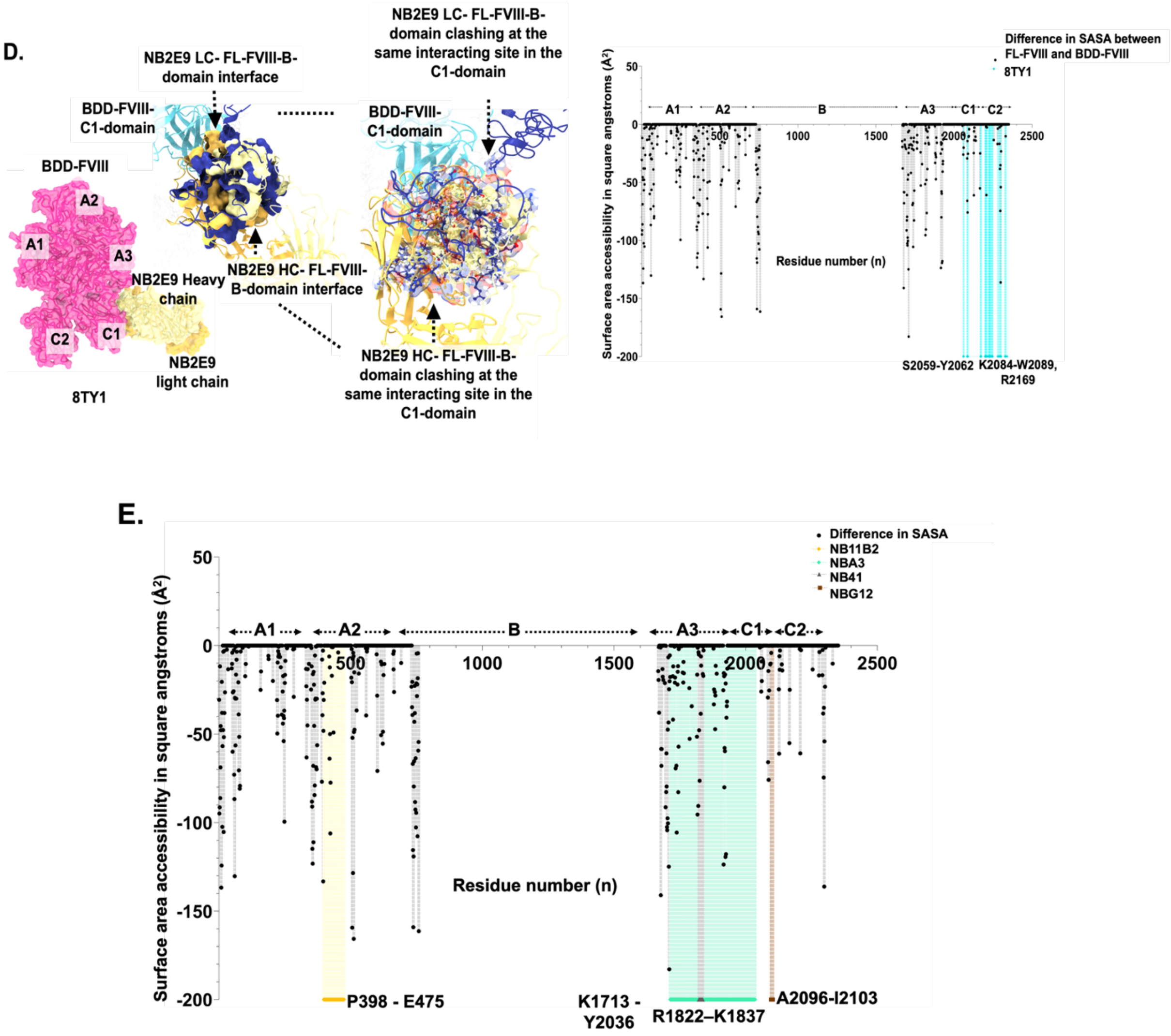


Figure 35. FVIII Surface area accessibility studies. A. 4KI5 - anti-C2 domain inhibitor containing G99 and 3E6 Fab Heavy and Light chains (HC and LC). The graph on the top right represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and BDD-FVIII (black, circles) and 4KI5 epitopes (orange, diamonds) and their respective comparative accessibility, which is depicted to be an average of -6.3 Å^2. The image on the top right depicts the 4KI5 binding residues being overlapped by the presence of the B-domain. **B. 4XZU -** anti-C2-domain inhibitor containing 3E6 Fab Heavy and Light chains (HC and LC). The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and BDD-FVIII (black, circles) and 4XZU epitopes (purple, downward triangles) and their respective comparative accessibility, which is depicted to be an average of -8 Å^2. The image on the top right depicts the 4XZU binding residues being overlapped by the presence of the B-domain. **C. 7K66 -** anti-C1-domain inhibitor containing the 2A9 Heavy and light chain (HC and LC). The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and BDD-FVIII (black, circles) and 7K66 epitopes (pink, circles) and their respective comparative accessibility, which is depicted to be an average of -10.89 Å^2. The image on the top right depicts the 7K66 binding residues being overlapped by the presence of the B-domain. **D. 8TY1 -** anti-C1-domain inhibitor containing the NB2E9 Heavy and light chain (HC and LC). The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and BDD-FVIII (black, circles) and 8TY1 epitopes

(yellow, squares) and their respective comparative accessibility, which is depicted to be an average of -24.21 Å^2. The image on the top right depicts the 8TY1 binding residues being overlapped by the presence of the B-domain. **E. NBAbs -** The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and BDD-FVIII (black, circles) and anti-C1 neutralizing antibody epitope - NBG12 (brown, square) -0.51 Å^2, anti-A3 neutralizing antibody epitope - NBA3 (mint green, circles), and NB41 (grey, triangles) -12.59 Å^2 and -6.75 Å^2 respectively and anti-A2 neutralizing antibody epitope: NB11B2 (violet, diamonds) -5.63 Å^2 shown to be influenced by the presence of the B-domain.

- **Inhibitory Antibodies:** Structural alignment of the *Gly/FC-FL-FVIII* with inhibitory antibody complexes revealed that the B-domain hinders the binding site epitopes. Any reduction in the FVIII solvent accessibility due to the presence of the B-domain will provoke a steric hindrance preventing inhibitors from reaching FVIII epitopes. The 4KI5 and 4XZU -anti-C2 (G99 and 3E6 Fab HC, LC) structure shows that B-domain residues (listed in Table 11) overlap with the G99 and 3E6 Fab HC, LC interaction site on the C2-domain, reducing its SASA by -6.3 Å^2 and -8 Å^2 respectively in *Gly/FC-FL-FVIII* (Figure 35). The 7K66-anti-C1 (2A9 HC and LC) structure revealed that B-domain regions (listed in Table 11) overlap with the interaction interfaces between 2A9 Fab HC, LC, and the C1-domain, reducing average accessibility by -10.9 Å^2 (Figure 35). Similarly, the 8TY1 - anti-C1 (NB2E9 Fab HC, LC) structure also exhibited B-domain residues (listed in Table 11) to overlap at the NB2E9 Fab HC, LC and the C1-domain interaction interface reducing its average SASA by -24.2 Å^2. Additionally, NBAbs - NBG12-anti-C1, NBA3, NB41-anti-A3 and NB11B2-anti-A2 epitopes displayed reduced SASA by -0.5 Å^2, -12.6 Å^2, -6.7 Å^2 and -5.6 Å^2 respectively (Figure 35). Most Ab interaction interfaces except NB33 and C2/G4k Fab (PDB-IDs: 8G6I and 1IQD) displayed significant B-domain overlap (Table 11).
    - **4KI5 -** anti-C2-domain inhibitor containing G99 and 3E6 Fab Heavy and Light chains (HC and LC). The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and 3CDZ (black, circles) and 4KI5 epitopes (purple, diamonds) and their respective comparative accessibility, which is depicted to be an average of -6.32 Å^2 (Figure 35A and Table 11).
    - **4XZU -** anti-C2-domain inhibitor containing 3E6 Fab Heavy and Light chains (HC and LC). The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and 3CDZ (black, circles) and 4XZU epitopes (green, diamonds) and their respective comparative accessibility, which is depicted to be an average of -8.07 Å^2 (Figure 35B and Table 11).
    - **7K66 -** anti-C1-domain inhibitor containing the 2A9 Heavy and light chain (HC and LC). The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and 3CDZ (black, circles) and 7K66 epitopes (light blue, diamonds) and their respective comparative accessibility, which is depicted to be an average of -10.89 Å^2 (Figure 35C and Table 11).

- **8TY1 -** anti-C1-domain inhibitor containing the NB2E9 Heavy and light chain (HC and LC). The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and 3CDZ (black, circles) and 7K66 epitopes (cyan, diamonds) and their respective comparative accessibility, which is depicted to be an average of -24.21 Å^2 (Figure 35D and Table 11).
- **NBAbs -** The graph represents the average difference in the accessibility values between *Gly/FC-FL-FVIII* and 3CDZ (black, circles) and anti-C1 neutralizing antibody epitope - NBG12 (brown, square) -0.51 Å^2, anti-A3 neutralizing antibody epitope - NBA3 (mint green, circles), and NB41 (grey, triangles) - 12.59 Å^2 and -6.75 Å^2 respectively and anti-A2 neutralizing antibody epitope: NB11B2 (yellow, diamond) -5.63 Å^2 shown to be influenced by the presence of the B-domain (Figure 35E and Table 11).

Inhibitor Abs (PDB-ID)	Epitopes	Average surface accessbility ($Å^2$)	Buried in the presence of B-domain	B-domain residues contacts	Functional region
G99 - Non-classical - C2-domain inhibitor PDB-ID: 4KI5	G2198-A2207 S2212-S2213 T2221-R2234 Q2241-W2247 L2280-S2282 W2290-L2292 K2298-V2301 R2326-Q2330, R2339	FL-FVIII : 46.67 $Å^2$ BDD-FVIII : 52.99 $Å^2$ Avg difference : -6.31 $Å^2$	Yes	K1492, E1495-G1509	Occlude residues implicated in vWF binding
3E6 - Classical - C2-domain inhibitor PDB-ID: 4XZU	G2198-D2207 T2210-S2212 S2223-R2234 Q2241-W2248, D2252 L2280-S2282 G2287-V2301 R2326-Q2330	FL-FVIII : 51.46 $Å^2$ BDD-FVIII : 59.53 $Å^2$ Avg difference : -8.07 $Å^2$	Yes	L1040-I1041 F1176-L1182 T1196-L1223	Regions involved in binding both Factors IXa and Xa
BO2C11 (G4k antibody) NB2C11- Classical - C2-domain inhibitor PDB-ID: 1IQD	F2215-A2220, G2233-R2239, Q2241-N2243, S2269-M2274, H2334-Q2335	FL-FVIII : 85.58 $Å^2$ BDD-FVIII : 85.58 $Å^2$ Avg difference : 0 $Å^2$	No	–	Interferes with the binding of factor VIII to phospholipid membrane surfaces and to von Willebrand factor - Neutralizing FVIII activity.
NB33- recombinant derivative of KM33 -	R2109-S2113, I2177-R2178	FL-FVIII : 51.23 $Å^2$ BDD-FVIII : 54.56 $Å^2$ Avg difference : -3.32 $Å^2$	Yes	–	Disrupts FVIII binding to lipoprotein receptor-related

Classical - C1-domain inhibitor PDB-ID: 8G6I					protein-1 (LRP1), which drives FVIII hepatic clearance and antigen presentation in dendritic cells - Neutralizing FVIII activity
2A9 - Classical C1-domain inhibitor PDB-ID: 7K66	F1762-Y1767 N1829, S2059, K2084-W2089, I2121-M2123, K2129-T2133, T2141-L2142, R2169-Y2175	FL-FVIII : 59.37 Å^2 BDD-FVIII : 70.26 Å^2 Avg difference : -10.89 Å^2	Yes	W1554, R1559-L1584, Y1590-F1592, E1603-D1615, L1620 N1626-A1628 I1632, E1640-K1647 R1650-R1667	Show interference with VWF binding and FVIII function.
NB2E9 - Non-classical C1-domain inhibitor PDB-ID: 8TY1	S2059-Y2062, K2084-W2089, R2169-H2174	FL-FVIII : 52.67 Å^2 BDD-FVIII : 76.88 Å^2 Avg difference : -24.21 Å^2	Yes	P1560-T1570 S1572-K1579 L1580-L1584 A1610-R1667	Neutralizing FVIII activity
NBG12 - non-classical C1-domain inhibitor	A2096-I2103	FL-FVIII : 16.86 Å^2 BDD-FVIII : 17.38 Å^2 Avg difference : -0.51 Å^2	Yes	–	Neutralizing FVIII activity
NB41- non-classical A3-domain inhibitor	R1822-K1837	FL-FVIII : 51.86 Å^2 BDD-FVIII : 64.46 Å^2 Avg difference : -12.59 Å^2	Yes	–	Anti-A3 inhibitors bind to an epitope that includes amino acids 1811-1818 and block the binding of FVIIIa to factor IXa -

					Neutralizing FVIII activity
NBA3 - Non-classical A3-domain inhibitor	K1713-Y2036	FL-FVIII : 27.43 Å^2 BDD-FVIII : 34.19 Å^2 Avg difference : -6.75 Å^2	Yes	–	Anti-A3 inhibitors bind to an epitope that includes amino acids 1811-1818 and block the binding of FVIIIa to factor IXa - Neutralizing FVIII activity
NB11B2 - non-classical A2-domain inhibitor	P398-E475	FL-FVIII : 40.35 Å^2 BDD-FVIII : 45.98 Å^2 Avg difference : -5.63 Å^2	Yes	–	Neutralizing FVIII activity

Table 11. Surface accessibility details between *Gly/FC-FL-FVIII* and BDD-FVIII for key epitope residues involved in binding to the inhibitory antibodies, demonstrating the impact of the B-domain and glycosylations.

3.5.2.B Surface area accessibility studies: FVIII Interacting partners

- **vWF** - The surface area accessibility graph on the right represents the reduced surface area accessibility in the binding regions due to the presence of the B-domain and the glycosylation. The graphs show the average difference in the accessibility values between *Gly/FC-FL-FVIII* and 3CDZ (black, circles). vWF interacting regions (navy blue) present in a3-acidic region and C1-domain displayed drastic reduction in the average accessibility -25.15 Å^2, -13.03 Å^2 respectively, while A3-domain binding residues displayed the average accessibility to -8.62 Å^2 and C2-domain residues showed -0.82 Å^2 which are depicted in the graph (Figure 36A, Table 12).
- **Thrombin**: The surface area accessibility graph on the top left shows the accessibility of thrombin cleavage sites (red, square) (R391, R759, and R1708) and thrombin binding regions (maroon, stars) (E340-E351 and E1680-D1684) to be reduced in *Gly/FC-FL-FVIII* than compared to 3CDZ. The difference between the thrombin cleavage sites being -76.93 Å^2, -161.34 Å^2, and -124.90 Å^2 respectively, and thrombin binding regions being -3.04 Å^2 and -27.92 Å^2 are depicted in the graph (Figure 36B, Table 12).

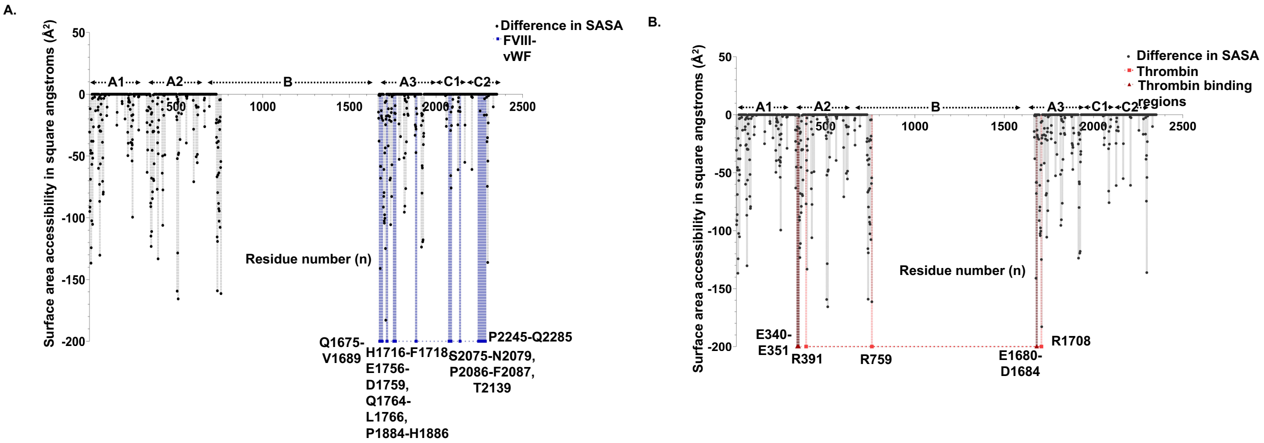


Figure 36. FVIII Surface area accessibility studies. A. vWF - The surface area accessibility graph on the right represents the reduced surface area accessibility in the binding regions due to the presence of the B-domain and the glycosylation. The graphs show the average difference in the accessibility values between *Gly/FC-FL-FVIII* and BDD-FVIII (black, circles). vWF interacting regions (blue, diamonds) present in a3-acidic region and C1-domain displayed drastic reduction in the average accessibility -25.15 Å^2, -13.03 Å^2 respectively, while A3-domain binding residues displayed the average accessibility to -8.62 Å^2 and C2-domain residues showed -0.82 Å^2 which are depicted in the graph. **B. Thrombin -** The surface area accessibility graph on the top left shows the accessibility of thrombin cleavage sites (red, square) (R391, R759 and R1708) and thrombin binding regions (maroon, triangles) (E340-E351 and E1680-D1684) to be reduced in *Gly/FC-FL-FVIII* than compared to BDD-FVIII. The difference between the thrombin cleavage sites being -76.93 Å^2, -161.34 Å^2, and -124.90 Å^2 respectively, and thrombin binding regions being -3.04 Å^2 and -27.92 Å^2 are depicted in the graph

Interacting regions/ Partners	**Domain**	**Residues**	**SASA - FL-FVIII (Å^2)**	**SASA - BDD-FVIII (Å^2)**	**Difference in SASA (Å^2)**
Thrombin cleavage sites	A1-, A2- and A3-domains	R391 R759 R1708	135.47 Å^2 71.56 Å^2 139.94 Å^2	212.40 Å^2 232.91 Å^2 264.84 Å^2	-76.93 Å^2 -161.34 Å^2 -124.90 Å^2
Thrombin binding sites	A1-domain	E340-E351	Avg difference : 54.64 Å^2	Avg difference : 57.68 Å^2	Avg difference : -3.04 Å^2
	A3-domain	E1680-D1684	Avg difference : 50.38 Å^2	Avg difference : 78.31 Å^2	Avg difference : -27.92 Å^2
vWF	a3-acidic region	Q1675-V1689	Avg difference : 57.69 Å^2	Avg difference : 82.84 Å^2	Avg difference : -25.15 Å^2
	A3-domain	H1716-F1718, E1756-D1759, Q1764-L1766, P1884-H1886	Avg difference : 53.67 Å^2	Avg difference : 62.30 Å^2	Avg difference : -8.62 Å^2
	C1-domain	S2075-N2079, P2086-F2087, T2139	Avg difference : 78.59 Å^2	Avg difference : 91.63 Å^2	Avg difference : -13.03 Å^2
	C2-domain	P2245-Q2285	Avg difference : 34.67 Å^2	Avg difference : 35.50 Å^2	Avg difference : -0.82 Å^2

Table 12. Surface accessibility details between *Gly/FC-FL-FVIII* and BDD-FVIII for key epitope residues involved in binding to the interacting partners - Thrombin cleavage and binding sites and vWF, demonstrating the impact of the B-domain and glycosylation.

3.5.2.C Surface area accessibility studies: FVIII Mutations

A total of 431 reported Hemophilic mutations within the FVIII-HC and LC were analyzed. Amongst these, 26 of 131 frameshifts (19.8 %), 17 of 70 nonsense (24.2 %), and 20 of 230 missense (8.5 %) mutations, that were reported to induce inhibitor formation, revealed reduced accessibility, i.e., SASA (average -41.9 Å^2) (Figure 37 and Figure 38). Among these, 35 mutations were located at the B-domain interaction interface (Figure 37, Figure 38A, Table 13). Additionally, 125 reported Hemophilic mutations within the B-domain were analyzed. Three of these mutations involved glycosylated residues (N919, N1020, N1461) and 28 were located at the B-domain interface (within FVIII) (Figure 37, Figure 38B, Table 14).

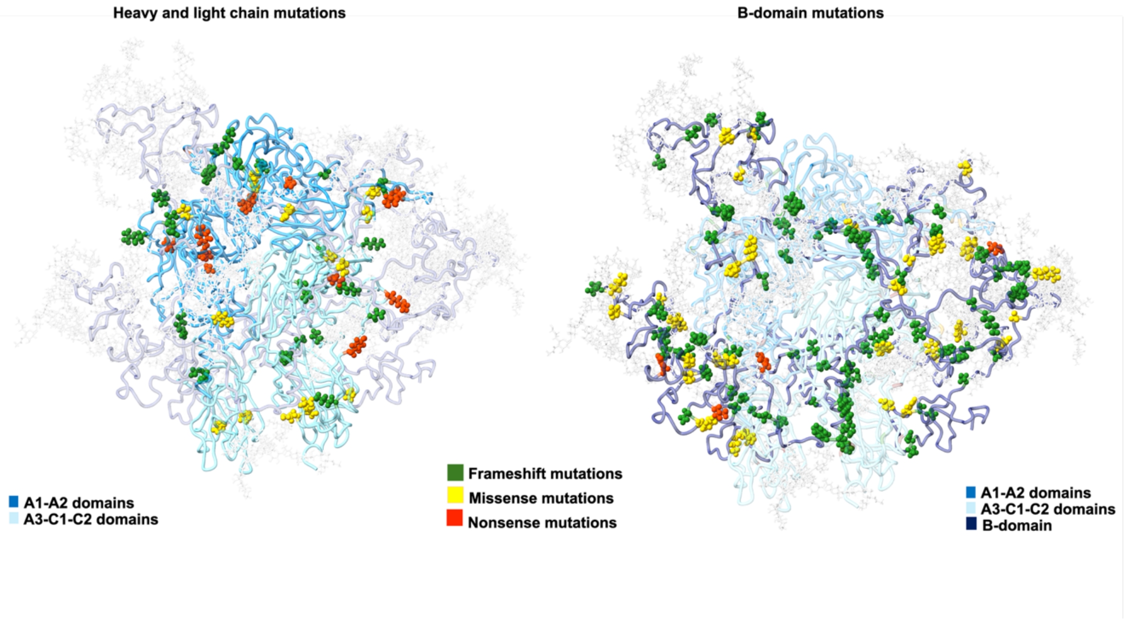


Figure 37. Mapping the Hemophilic mutations on the *Gly/FC-FL-FVIII* model. Structural Mapping of interface residues on heavy and light chains that are clinically reported to show frameshift (green)/missense (yellow)/nonsense (red) mutations and show variability in SASA in the presence of B-domain. Additionally, all B-domain clinically reported frameshift (green)/missense (yellow)/nonsense (red) mutations were also mapped on the *Gly/FC-FL-FVIII* model.

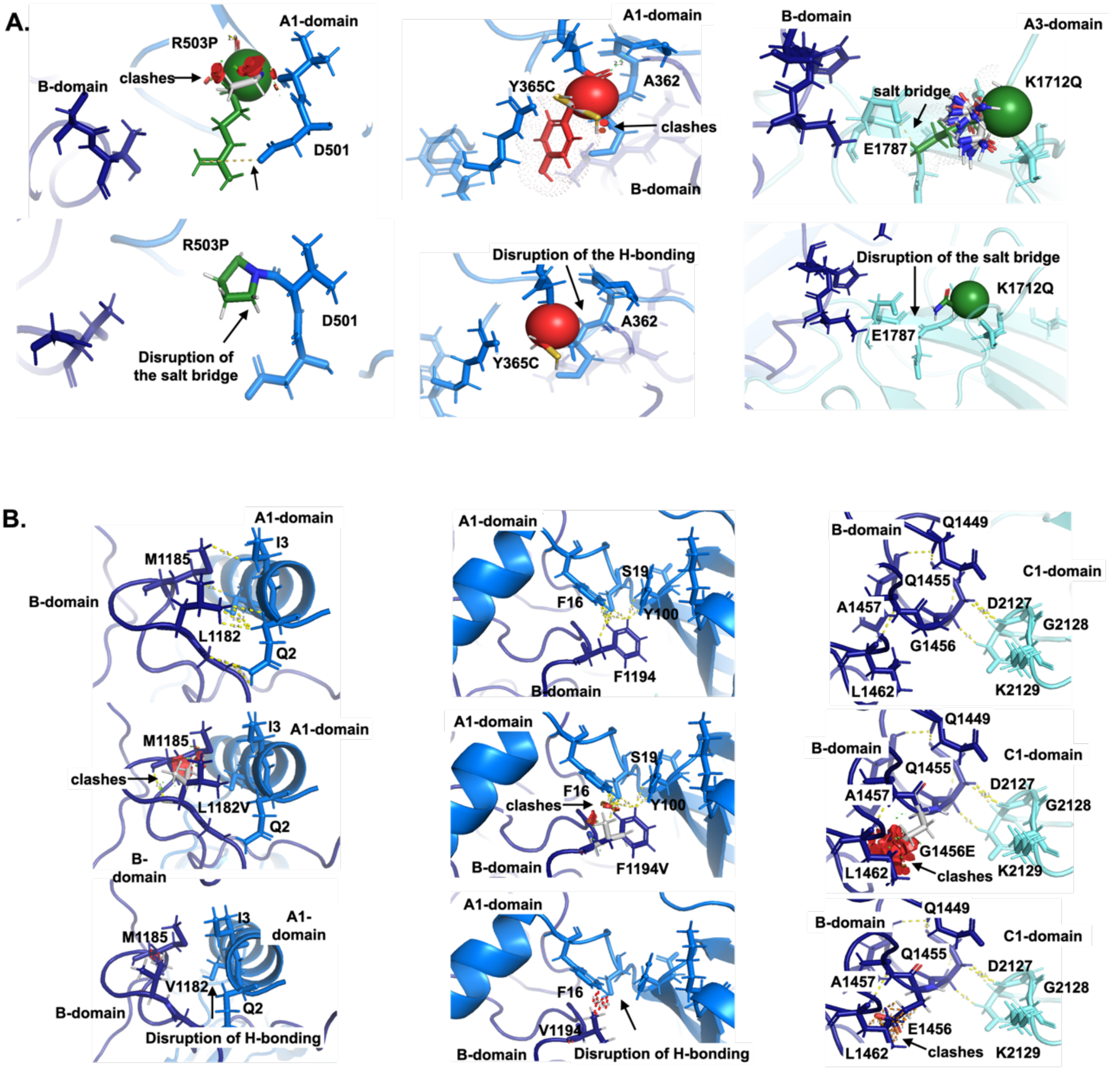


Figure 38. Mapping the Hemophilic mutations on the model. A. HC and LC mutations found in the A1- R503P (left), Y365C (middle), and A3-domains K1712Q (right) were found to impose steric clashes and disruption of salt bridges and H-bonding. **B.** B-domain missense mutations found at the interfaces of B-domain and A1-domain L1182V (left) and F1194V (center) disrupted the H-bonding between them and forming steric clashes. Similarly, the missense mutation of G1437Q (right) was found to impose steric clashes with the surrounding residues.

Mature protein	Mutated protein	Mutation Type	Mechanism	Domain	In Poly A	Reported Severity	History of Inhibitors	SASA FL-FVIII (Å^2)	SASA BDD-FVIII (Å^2)	Difference in SASA (Å^2)	Buried in the presence of B-domain	Interface residue	Functional relevance	Reference
p.(Ser6Leufs*14)	Ser-13Leufs*14	Frameshift	Duplication	Signal	N	Severe	Yes	0	68.844	**-68.844**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34. E2382–E2392 (2013)
p.(Arg23Valfs*18)	Arg4Valfs*18	Frameshift	Deletion/Insertion	A1	N	Severe	Not Reported	109.793	136.22	**-26.427**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Phe59Serfs*13)	Phe40Serfs*13	Frameshift	Deletion	A1	N	Severe	Yes	90.56	177.273	**-86.713**	Yes	**Y**	**Disruption of Salt bridge**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Val64Met	Val64Met	Missense	Substitution	A1	N	Mild	Yes	30.784	31.518	**-0.734**	Yes	**N**		Repesse Y, Slaoui M, Ferrandiz D, et al, Journal of Thrombosis & Haemostasis 5:1469-76, 17445092
p.Tyr100*	Tyr81*	Nonsense	Substitution	A1	N	Severe	Yes	45.309	62.636	**-17.327**	Yes	**Y**		Wang XF, Zhao YQ, Yang RC, et al, Haemophilia 16:632-9, 20331753
p.Glu223Lys	Glu204Lys	Missense	Substitution	A1	N	Severe	Yes	1.735	11.318	**-9.583**	Yes	**N**		Williams IJ, Abuzenadah A, Winship PR, et al, Thrombosis & Haemostasis 79:723-6, 9569180
p.Ser226Ile	Ser226Ile	Missense	Substitution	A1	N	Mild/Severe	Yes	0	6.238	**-6.238**	Yes	**N**		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011, http://hadb.org.uk
p.Arg355Tyrfs*3	Arg336Tyrfs*3	Frameshift	Duplication	A1	N	Severe	Yes	135.328	250.204	**-114.876**	Yes	**N**	**Located near R391 thrombin cleavage site and E351 - thrombin binding sites**	Ravanbod S, Rassoulzadegan M, Rastegar-Lari G, et al, Haemophilia 18:e340-6, 22117735
p.(Asn358Ilefs*60)	Asn339Ilefs*60	Frameshift	Deletion	a1	N	Severe	Not Reported	23.776	91.704	**-67.928**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Asn359*	Asn340*	Frameshift	Duplication	a1	N	Severe	Yes	16.413	127.546	**-111.133**	Yes	**Y**	**Located near R759 thrombin cleavage site.**	HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011, http://hadb.org.uk
p.(Glu361*)	Glu342*	Nonsense	Substitution	a1	N	Severe	Yes	90.735	112.842	**-22.107**	Yes	**Y**		Konkle BA, Fletcher SN, Johnsen JM, et al, My Life Our Future, unpublished,
p.Tyr365Cys	Tyr365Cys	Missense	Substitution	A1	N	Mild	Yes	96.942	102.52	**-5.578**	Yes	**Y**	**Tyrosine sulfation site**	HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011, http://hadb.org.uk
p.(Tyr365*)	Tyr346*	Nonsense	Substitution	a1	N	Severe	Yes	96.942	102.52	**-5.578**	Yes	**Y**		Eckhardt CL, van Velzen AS, Peters M, et al, Blood 122:1954-62, 23926300
p.(Thr400Ilefs*18)	p.Thr381Ilefs*18	Frameshift	Deletion	A2	N	Severe	Not reported	61.286	82.29	**-21.004**	Yes	**N**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Ser428Trpfs*7)	Ser409Trpfs*7	Frameshift	Deletion	A2	N	Severe	Not reported	0.555	1.316	**-0.761**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Arg503Profs*28)	Arg484Profs*28	Frameshift	Deletion	A2	N	Moderate	Yes	52.049	70.158	**-18.109**	Yes	**Y**	**Disruption of Salt bridge**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Gly513Ser	Gly494Ser	Missense	Substitution	A2	N	Moderate	Yes	43.683	92.48	**-48.797**	Yes	**N**		Liu ML, Nakaya S and Thompson AR, Thrombosis & Haemostasis 87:273-6, 11858487
p.Pro521Glnfs*14	Pro502Glnfs*14	Frameshift	Deletion	A2	N	Severe	Yes	46.299	46.452	**-0.153**	Yes	**N**		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011, http://hadb.org.uk
p.(Lys529Asnfs*6)	Lys510Asnfs*6	Frameshift	Deletion	A2	N	Severe	Yes	30.15	38.929	**-8.779**	Yes	**Y**	**Disruption of Salt bridge**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
Arg531Cys	Arg531Cys	Missense	Substitution	A1	N	Non-Severe	Yes	42.376	58.201	**-15.825**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Arg602*	Arg583*	Nonsense	Substitution	A2	N	Moderate/Severe	Yes	19.7	90.545	**-70.845**	Yes	**N**	**Disruption of Salt bridge**	HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011, http://hadb.org.uk
p.Leu622Profs*3	Leu603Profs*3	Frameshift	Duplication	A2	N	Severe	Yes	79.945	135.337	**-55.392**	Yes	**N**		Kim HJ, Chung HS, Kim SK, et al, Haemophilia 18:1008-13, 22741565
Asn618Ser	Asn618Ser	Missense	Substitution	A2	N	Non-Severe	Yes	91.113	101.744	**-10.631**	Yes	**N**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Tyr655X	Tyr655X	Non missense mutation	Substitution	A1	N	Severe	No	1.346	17.811	**-16.465**	Yes	**N**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Tyr737Serfs*2	Tyr718Serfs*2	Frameshift	Deletion/Insertion	a2	N	Severe	Yes	6.888	122.436	**-115.548**	Yes	**N**	**Located near R759 thrombin cleavage site and Tyrosine sulfation site**	Bogdanova N, Markoff A, Pollmann H, et al, Human Mutation 20:236-7, 12204009
p.Glu739Lys*	Glu739Lys*	Missense	Substitution	A1	N	Non-Severe	_	3.988	123.115	**-119.127**	Yes	**Y**	**Located near R759 thrombin cleavage site.**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Leu750Val	Leu731Val	Missense	Substitution	a2	N	Moderate	Yes	63.474	84.794	**-21.32**	Yes	**Y**	**Located near R759 thrombin cleavage site.**	Green PM, Bagnall RD, Waseem NH, et al, British Journal of Haematology 143:115-28, 18691168
p.(Gln1678*)	Gln1659*	Nonsense	Substitution	a3	N	Severe	Yes	33.064	174.134	**-141.07**	Yes	**Y**	**loacated in the a3 acidic region - vWF binding region and Disruption of H bonding**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Glu1679*)	Glu1660*	Nonsense	Substitution	a3	N	Severe	Yes	52.635	111.065	**-58.43**	Yes	**Y**	**loacated in the a3 acidic region - vWF binding region**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Glu1680*	Glu1661*	Nonsense	Substitution	a3	N	Severe	Yes	30.642	98.597	**-67.955**	Yes	**Y**	**D1684 thrombin binding site and loacated in the a3 acidic region - vWF binding region**	Guillet B, Lambert T, d'Oiron R, et al, Human Mutation 27:676-85, 16786531
p.Lys1693Argfs*38	Lys1674Argfs*38	Frameshift	Deletion	a3	Y	Severe	Yes	107.068	128.261	**-21.193**	Yes	**Y**	**Loacated in the vWF binding region and Disruption of Salt bridge**	Kim HJ. Chung HS, Kim SK, et al, Haemophilia 18:1008-13, 22741565
p.Tyr1699Phe	Tyr1699Phe	Missense	Substitution	A3	N	Non-Severe	_	98.366	159.256	**-60.89**	Yes	**N**	**Direct effect on binding of FVIII to vWF and Tyrosine sulfation site**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34. E2382–E2392 (2013)
p.(Tyr1699*)	Tyr1680*	Nonsense	Substitution	a3	N	Severe	Yes	98.366	159.256	**-60.89**	Yes	**N**		Konkle BA, Fletcher SN, Johnsen JM, et al, My Life Our Future, unpublished
p.Asn1704Lysfs*28	Asn1685Lysfs*28	Frameshift	Duplication	a3	Y	Severe	Yes	41.898	89.375	**-47.477**	Yes	**N**	**Located near R1708 thrombin cleavage site.**	Miller CH, Benson J, Ellingsen D, et al, Haemophilia 18:375-82, 22103590
p.Arg1708His*	Arg1708His*	Missense	Substitution	A3	N	Non-Severe	_	139.945	264.848	**-124.903**	Yes	**N**	**Located within thrombin cleavage site.**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Lys1712Asnfs*19)	Lys1693Asnfs*19	Frameshift	Deletion	A3	N	Severe	Yes	74.251	79.782	**-5.531**	Yes	**Y**	**Disruption of Salt bridge**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Ser1743Trpfs*16)	Gln1723Glnfs*17	Frameshift	Deletion	A3	N	Severe	Not reported	19.087	30.74	**-11.653**	Yes	**Y**		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011, http://hadb.org.uk
p.(Val1746delfs*)	Val1727delfs*	Frameshift	Deletion/Insertion	A3	N	Severe	Not reported	29.392	31.387	**-1.995**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Glu1787Lys)	Glu1768Lys	Missense	Substitution	A3	N	Moderate	Not reported	31.034	83.685	**-52.651**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Gln1815*	Gln1796*	Nonsense	Substitution	A3	N	Severe	Yes	92.363	143.884	**-51.521**	Yes	**Y**	**Disruption of H bonding**	Lin SW, Lin SR and Shen MC, Genomics 18:496-504, 8307558
p.Gln1817*	Gln1798*	Nonsense	Substitution	A3	N	Severe	Yes	110.404	205.856	**-95.452**	Yes	**N**		Pinto P, Ghosh K and Shetty S, Mutation Research 786:27-33, 26897466
p.(Gln1817Glufs*55)	Gln1798Glufs*55	Frameshift	Duplication	A3	N	Severe	Yes	110.404	205.856	**-95.452**	Yes	**N**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Gly1818X	Gly1818X	Nonmissense mutation	Substitution	A3	N	Severe	No	36.542	37.741	**-1.199**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Lys1827Serfs*44)	Lys1808Serfs*44	Frameshift	Deletion	A3	N	Severe	Not reported	120.151	120.211	**-0.06**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Lys1837*	Lys1818*	Nonsense	Substitution	A3	N	Severe	Yes	69.619	85.707	**-16.088**	Yes	**Y**	**Disruption of Salt bridge**	Konkle BA, Fletcher SN, Johnsen JM, et al, My Life Our Future, unpublished
p.Glu1915*	Glu1896*	Nonsense	Substitution	A3	N	Severe	Yes	82.296	107.541	**-25.245**	Yes	**Y**		Albanez S, Ruiz-Saez A, Boadas A, et al, Haemophilia 17:e913-8, 21371196
p.(Cys1922Alafs*23)	Cys1903Alafs*23	Frameshift	Deletion	A3	N	Severe	Not reported	11.512	27.905	**-16.393**	Yes	**Y**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Glu2085*	Glu2066*	Nonsense	Substitution	C1	N	Severe	Yes	7.802	73.726	**-65.924**	Yes	**Y**	**Loacated in the C1-vWF binding region**	Liu ML, Nakaya S and Thompson AR, Thrombosis & Haemostasis 87:273-6, 11858487
p.(Glu2085Glyfs*14)	Glu2066Glyfs*14	Frameshift	Deletion	C1	N	Severe	Not reported	7.802	73.726	**-65.924**	Yes	**Y**	**7K66 inhibitor binding residue**	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Asp2127Valfs*16	Asp2108Valfs*16	Frameshift	Deletion	C1	N	Severe	Yes	6.618	67.695	**-61.077**	Yes	**Y**		Chuansumrit A, Sasanakul W, Sirachainan N, et al, Haemophilia 23:e518-23, 28891589
p.(Asn2137Lysfs*24)	Asn2118Lysfs*24	Frameshift	Duplication	C1	N	Severe	Yes	40.559	51.222	**-10.663**	Yes	**N**		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Arg2166Hisfs*20	Arg2147Hisfs*20	Frameshift	Deletion	C1	N	Severe	Yes	16.239	71.37	**-55.131**	Yes	**Y**	**Disruption of Salt bridge**	Vinciguerra C, Zawadzki C, Dargaud Y, et al, Thrombosis & Haemostasis 95:593-9, 16601827
p.Gln2254Argfs*4	Gln2235Argfs*4	Frameshift	Deletion	C2	N	Severe	Yes	91.433	104.921	**-13.488**	Yes	**Y**	**Loacated in the vWF binding region**	Guillet B. Lambert T, d'Oiron R, et al, Human Mutation 27:676-85, 16786531
p.Gln2303*	Gln2284*	Nonsense	Substitution	C2	N	Severe	Yes	88.07	89.205	**-1.135**	Yes	**N**		Beskorovainaya TS, MilovidovaTB, Schagina OA, et al, Russian Journal of Genetics 55:1015-24, https://doi.org/10.1134/S1022795419080027
p.Gln2330*	Gln2311*	Nonsense	Substitution	C2	N	Severe	Yes	80.554	90.738	**-10.184**	Yes	**Y**		Goodeve AC, Williams I, Bray GL, et al, Thrombosis & Haemostasis 83:844-8, 10896236

Table 13. Hemophilia mutations. Comparison of surface accessibility between *Gly/FC-FL-FVIII* and BDD-FVIII for residues involved in mutations (A1-, A2-, A3-, C1- and C2-domains) causing severe/moderate Hemophilia-A and inhibitor formations mutations based on accessibility changes that affect binding sites and overall protein function.

Mature protein	Mutated protein	Mutation Type	Mechanism	Domain	In Poly A	Reported Severity	History of Inhibitors	SASA FL-FVIII (Å)	Interface residue	Functional relevance	Reference
p.Asn753Lysfs*33	Asn734Lysfs*33	Frameshift	Deletion	B	Y	Severe	Yes	27.339	N		Rydz N, Leggo J, Tinlin S, et al, American Journal of Hematology 88:1030-4, 23913812
p.Pro758Alafs*12	Pro739Alafs*12	Frameshift	Insertion	B	N	Severe	Yes	29.41	N		Rydz N. Leggo J. Tinlin S. et al. American Journal of Hematology 88:1030-4. 23913812
p.(Arg759*)	Arg740*	Nonsense	Substitution	B	N	Severe	Not reported	71.566	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34. E2382–E2392 (2013)
p.Ser765*	Ser746*	Nonsense	Substitution	B	N	Severe	Yes	59.857	N		Miller CH. Benson J. Ellingsen D. et al. HIRS, unpublished, CHAMP
p.(Glu782*)	Glu763*	Nonsense	Substitution	B	N	Severe	Yes	119.119	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Arg795*	Arg776*	Nonsense	Substitution	B	N	Not reported	Yes	95.633	N		Zahari M, Sulaiman SA, Othman Z, et al, Mediterranean Journal of Hematology and Infectious Diseases 10:e2018056, 30210749
p.(Met798Thr)	Met779Thr	Missense	Substitution	B	N	Severe	Yes	71.601	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Asn803Lysfs*8	Asn784Lysfs*8	Frameshift	Deletion	B	N	Severe	Yes	23.924	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Ser807Lysfs*2	Ser788Lysfs*2	Frameshift	Duplication	B	N	Severe	Yes	36.349	N		Green PM, Bagnall RD, Waseem NH, et al, British Journal of Haematology 143:115-28, 18691168
p.(Leu813Serfs*10)	Leu794Serfs*10	Frameshift	Deletion	B	N	Severe	Not reported	109.319	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Arg814*	Arg795*	Nonsense	Substitution	B	N	Severe	Yes	42.425	N		Pattinson JK, Millar DS, McVey JH, et al, Pattinson JK, Millar DS, McVey JH, et al, 1979502
p.Thr818Aspfs*9	Thr799Aspfs*9	Frameshift	Duplication	B	N	Severe	Yes	48.927	N		Guo Z, Yang L, Qin X, et al, Clinical and Applied Thrombosis/Hemostasis 24:70-78, 28056528
p.(Leu822Tyrfs*23)	Leu803Tyrfs*23	Frameshift	Deletion	B	N	Severe	Yes	121.982	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Ala830Profs*15	Ala811Profs*15	Frameshift	Deletion	B	N	Severe	Yes	48.758	N		Payne AB, Miller CH, Hemophilia Inhibitor Research Study Investigators, CHAMP, unpublished, CHAMP
p.Lys831Asnfs*14	Lys812Asnfs*14	Frameshift	Deletion	B	N	Severe	Yes	126.32	N		Rossetti LC, Radic CP, Candela M, et al, Haematologica 92:842-5, 17550859
p.Pro841Leufs*4	Pro822Leufs*4	Frameshift	Deletion	B	N	Moderate	Yes	37.196	N		Miller CH, Benson J, Ellingsen D, et al, HIRS, unpublished, CHAMP
p.(His855Leufs*16)	His836Leufs*16	Frameshift	Deletion	B	N	Severe	Yes	79.479	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Gln859*	Gln840*	Nonsense	Substitution	B	N	Severe	Yes	13.837	N		Wang XF, Zhao YQ, Yang RC, et al, Haemophilia 16:632-9, 20331753
p.Asp865Thrfs*12	Asp846Thrfs*12	Frameshift	Deletion	B	N	Severe	Yes	56.492	N		David D, Ventura C, Moreira I, et al, Haematologica 91:840-3, 16769589
p.Glu871Serfs*6	Glu852Serfs*6	Frameshift	Deletion	B	N	Severe	Yes	142.828	N		Bogdanova N, Markoff A, Pollmann H, et al, Bogdanova N, Markoff A, Pollmann H, et al, 12204009
p.Ser872*	Ser853*	Nonsense	Substitution	B	N	Severe	Yes	60.009	N		David D, Ventura C, Moreira I, et al, Haematologica 91:840-3, 16769589
p.(Leu874Profs*3)	Leu855Profs*3	Frameshift	Deletion	B	N	Severe	Not reported	159.671	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Lys881Thrfs*13)	Lys6862Thrfs*13	Frameshift	Deletion	B	N	Severe	Yes	98.218	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Leu882Argfs*9	Leu863Argfs*9	Frameshift	Deletion	B	N	Severe	Yes	64.594	Y		Nair PS, Shetty SD, Chandrakala S, et al, PLoS ONE 9 (5): e97337, 24845853
p.Gly883Leufs*8	Gly864Leufs*8	Frameshift	Deletion/Insertion	B	N	Severe	Yes	65.552	Y		Nakaya S, Liu ML and Thompson AR, British Journal of Haematology 115:977-82, 11843836
p.Ser901*	Ser882*	Nonsense	Substitution	B	N	Severe	Yes	74.829	N		Konkle BA, Fletcher SN, Johnsen JM, et al, My Life Our Future, unpublished,
p.(Ser906*)	Ser887*	Nonsense	Insertion	B	N	Severe	Not reported	40.14	N		Bai N, Zhu X, Zhao Z, et al, Blood Coagulation and Fibrinolysis 28:509-13, 28252515
p.(C390Asn912Phefs*6)	Asn893Phefs*6	Frameshift	Deletion	B	N	Severe	Not reported	50.404	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Asn919Lysfs*5)	Asn900Lysfs*5	Frameshift	Deletion	B	N	Severe	Not reported	132.578	N	Glycosylation site	CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Leu923Phefs*11	Leu904Phefs*11	Frameshift	Insertion	B	N	Severe	Yes	116.096	N		Chuansumrit A, Sasanakul W, Sirachainan N, et al, Haemophilia 23:e518-23, 28891589
p.Leu923*	Leu904*	Nonsense	Substitution	B	N	Severe	Yes	116.096	N		Vencesla A, Corral-Rodriguez MA, Baena M, et al, Blood 111:3468-78,
p.(Asn961Ilefs*7)	Asn942Ilefs*7	Frameshift	Deletion	B	N	Severe	Not reported	1.35	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Asn982Lysfs*9	Asn963Lysfs*9	Frameshift	Duplication	B	Y	Moderate/Severe	Yes	66.275	N		Lin SY, Su YN, Hung CC, et al, BMC Medical Genetics 9:53, 18565236
p.Glu987Argfs*17	Glu968Argfs*17	Frameshift	Deletion	B	N	Severe	Yes	102.679	N		Nakaya S, Liu ML and Thompson AR, British Journal of Haematology 115:977-82, 11843836
p.(Ala1001Leufs*3)	Ala982Leufs*3	Frameshift	Deletion	B	N	Severe	Not reported	76.247	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Leu1002Phefs*2	Leu983Phefs*2	Frameshift	Deletion	B	N	Severe	Yes	98.379	Y		Albanez S, Ruiz-Saez A, Boadas A, et al, Haemophilia 17:e913-8, 21371196
p.(C2492Thr1004Leufs*14)	Thr985Leufs*14	Frameshift	Deletion	B	N	Severe	Not reported	78.442	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Asn1007Lysfs*7)	Asn988Lysfs*7	Frameshift	Insertion	B	N	Severe	Not reported	68.09	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Leu1017*	Leu998*	Nonsense	Substitution	B	N	Severe	Yes	89.394	Y		Miller CH, Benson J, Ellingsen D, et al, Haemophilia 18:375-82, 22103590
p.(Lys1018*)	Lys999*	Nonsense	Substitution	B	Y	Severe	Yes	99.944	Y		Konkle BA, Fletcher SM, Johnsen JM, et al, My Life Our Future, unpublished,
p.(Asn1020Lysfs*20)	Asn1001Lysfs*20	Frameshift	Deletion	B	N	Severe	Yes	86.064	N	Glycosylation site	HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Lys1031Leufs*9	Lys1012Leufs*9	Frameshift	Deletion	B	Y	Severe	Yes	83.423	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.(Trp1048*)	Trp1029*	Nonsense	Substitution	B	N	Severe	Not reported	59.664	Y		Leuer M, Oldenburg J, Lavergne JM, et al, Leuer M, Oldenburg J, Lavergne JM, et al, 11410838
p.Ile1051Serfs*3	Ile1032Serfs*3	Frameshift	Insertion	B	N	Severe	Yes	12.532	Y		Nakaya S, Liu ML and Thompson AR, British Journal of Haematology 115:977-82, 11843836
p.Ile1051Argfs*3	Ile1032Argfs*3	Frameshift	Deletion	B	N	Not reported	Yes	12.532	N		Rosset C, Gorziza RP, Botton MR, et al, Blood Coagulation & Fibrinolysis 25:125-7, 23963097
p.(Glu1057*)	Glu1038*	Nonsense	Duplication	B	N	Moderate	Yes	22.063	N		He Z, Chen S and Jiang W, Blood Cells Molecules & Diseases 44:124-5, 20018532
p.Lys1059*	Lys1040*	Nonsense	Substitution	B	N	Severe	Yes	129.868	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.His1066Tyr	His1047Tyr	Missense	Substitution	B	N	Severe	Yes	0.406	N		Habart D, Kalabova D, Novotny M, et al, Journal of Thrombosis & Haemostasis 1:773-81, 12871415
p.(Asn1092Profs*46)	Asn1073Profs*46	Frameshift	Deletion/Insertion	B	N	Severe	Yes	10.834	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Gln1098*	Gln1079*	Nonsense	Substitution	B	N	Severe	Yes	12.187	N		Ivaskevicius V, Jurgutis R, Rost S, et al, Ivaskevicius V, Jurgutis R, Rost S, et al, 11298607
p.Lys1100*	Lys1081*	Nonsense	Substitution	B	N	Severe	Yes	104.266	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Glu1101Argfs*17	Glu1082Argfs*17	Frameshift	Duplication	B	Y	Moderate/Severe	Yes	97.073	N		0, 0, 7794769
p.Phe1120Ilefs*46	Phe1101Ilefs*46	Frameshift	Duplication	B	N	Severe	Yes	66.565	N		Cutler JA, Mitchell MJ, Smith MP, et al, Cutler JA, Mitchell MJ, Smith MP, et al, 11857744
p.Ser1124*	Ser1105*	Nonsense	Substitution	B	N	Severe	Yes	101.653	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Trp1127*	Trp1108*	Nonsense	Substitution	B	N	Severe	Yes	202.933	Y		Miller CH, Benson J, Ellingsen D, et al, HIRS, unpublished, CHAMP
p.Gln1129Lysfs*9	Gln1110Lysfs*9	Frameshift	Deletion	B	N	Severe	Yes	102.771	N		Bogdanova N, Markoff A, Pollmann H, et al, Human Mutation 26:249-54, 16086318
p.(Asn1135Glu*31)	Asn1114Glu*31	Frameshift	Duplication	B	N	Moderate	Not reported	22.781	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Ser1136Leufs*2)	Ser1117Leufs*2	Frameshift	Deletion	B	N	Severe	Yes	49.665	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Leu1137Glnfs*28	Leu1118Glnfs*28	Frameshift	Deletion	B	N	Severe	Yes	37.526	Y		Gouw SC, Van Der Bom JG, Van Den Berg HM, et al, Haemophilia 17:275-81, 21070499
p.(Asn1138Thrfs*11)	Asn1119Thrfs*11	Frameshift	Deletion	B	N	Severe	Not reported	31.03	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Pro1145Leufs*15)	Pro1126Leufs*15	Frameshift	Deletion	B	N	Severe	Not reported	114.478	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Lys1166*	Lys1147*	Nonsense	Substitution	B	N	Severe	Yes	103.736	N		Vinciguerra C, Zawadzki C, Dargaud Y, et al, Thrombosis & Haemostasis 95:593-9, 16601827
p.Asn1167Lysfs*9	Asn1148Lysfs*9	Frameshift	Duplication	B	Y	Severe	Yes	38.013	N		Rossetti LC, Radic CP, Candela M, et al, Haematologica 92:842-5, 17550859
p.Asn1167Lysfs*3	Asn1148Lysfs*3	Frameshift	Deletion	B	Y	Severe	Yes	38.013	N		Miller CH, Benson J, Ellingsen D, et al, HIRS, unpublished, CHAMP
p.Asp1179Glyfs*19	Asp1160Glyfs*19	Frameshift	Duplication	B	N	Severe	Yes	61.99	Y		Pinto P, Ghosh K and Shetty S, Mutation Research 786:27-33, 26897466
p.Leu1182Val	Leu1163Val	Missense	Substitution	B	N	Severe	Yes	63.099	Y		Payne AB, Miller CH, Hemophilia Inhibitor Research Study Investigators, CHAMP, unpublished, CHAMP
p.Ser1189Lysfs*9	Ser1170Lysfs*9	Frameshift	Duplication	B	Y	Severe	Yes	6.621	Y		Nakaya S, Liu ML and Thompson AR, British Journal of Haematology 115:977-82, 11843836
p.Phe1194Val	Phe1175Val	Missense	Substitution	B	N	Mild	Yes	71.202	Y		Mauser-Bunschoten EP, Den Uijl IE, Schutgens RE, et al, Haemophilia 18:263-7, 21851507
p.(Asp1199Glufs*19)	Asp1180Glufs*19	Frameshift	Deletion	B	N	Severe	Yes	44.07	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Lys1211Asnfs*30	Lys1192Asnfs*30	Frameshift	Insertion	B	Y	Severe	Yes	154.061	Y		Debeljak M, Kitanovski L, Trampus Bakija A, et al, Haemophilia 18:e420-3, 22958177
p.Lys1212Asnfs*28	Lys1193Asnfs*28	Frameshift	Deletion	B	Y	Severe	Yes	133.574	Y		Gouw SC, Van Der Bom JG, Van Den Berg HM, et al, Haemophilia 17:275-81, 21070499
p.Ile1213Phefs*5	Ile1194Phefs*5	Frameshift	Deletion	B	Y	Moderate/Severe	Yes	58.585	Y		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Lys1219Argfs*5	Lys1200Argfs*5	Frameshift	Deletion	B	Y	Severe	Yes	43.609	N		Rydz N, Leggo J, Tinlin S, et al, American Journal of Hematology 88:1030-4, 23913812
p.Ile1224Thrfs*5	Ile1205Thrfs*5	Frameshift	Deletion	B	N	Moderate	Yes	83.338	N		Wang XF, Zhao YQ, Yang RC, et al, Haemophilia 16:632-9, 20331753
p.Thr1235Serfs*5	Thr1216Serfs*5	Frameshift	Deletion	B	N	Severe	Yes	56.17	N		Miller CH, Benson J, Ellingsen D, et al, HIRS, unpublished, CHAMP
p.Arg1252Serfs*4	Arg1233Serfs*4	Frameshift	Deletion	B	N	Severe	Yes	99.608	N		Rossetti LC, Radic CP, Candela M, et al, Haematologica 92:842-5, 17550859
p.(Ala1264Valfs*10)	Ala1245Valfs*10	Frameshift	Deletion	B	N	Severe	Yes	67.856	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Ser1272*	Ser1253*	Nonsense	Substitution	B	N	Severe	Yes	39.992	N		Payne AB, Miller CH, Hemophilia Inhibitor Research Study Investigators, CHAMP, unpublished, CHAMP
p.Gly1291Argfs*29	Gly1272Argfs*29	Frameshift	Duplication	B	Y	Severe	Yes	49.999	N		Green PM, Bagnall RD, Waseem NH, et al, British Journal of Haematology 143:115-28, 18691168
p.Gln1302*	Gln1283*	Nonsense	Substitution	B	N	Severe	Yes	95.528	Y		Guillet B, Lambert T, d'Oiron R, et al, Human Mutation 27:676-85, 16786531
p.(Ala1311Phefs*9)	Ala1292Phefs*9	Frameshift	Deletion/Insertion	B	N	Severe	Not reported	31.423	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Arg1332Leufs*2	Arg1313Leufs*2	Frameshift	Deletion	B	N	Severe	Yes	16.993	N		Vidal F, Farssac E, Altisent C, et al, Vidal F, Farssac E, Altisent C, et al, 11341489
p.(Arg1338Lysfs*10)	Arg1319Lysfs*10	Frameshift	Insertion	B	N	Severe	Yes	71.309	Y		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Arg1349Lysfs*6	Arg1330Lysfs*6	Frameshift	Duplication	B	Y	Severe	Yes	42.046	N		Ravanbod S, Rassoulzadegan M, Rastegar-Lari G, et al, Haemophilia 18:e340-6, 22117735
p.Tyr1376*	Tyr1357*	Frameshift	Deletion	B	N	Severe	Yes	143.667	N		Konkle BA, Fletcher SM, Johnsen JM, et al, My Life Our Future, unpublished,
p.Ala1383Profs*41	Ala1362Profs*41	Frameshift	Deletion	B	Y	Not reported	Yes	62.793	N		Miller CH, Benson J, Ellingsen D, et al, HIRS, unpublished, CHAMP
p.(Gln1386Asnfs*52)	Gln1367Asn*52	Frameshift	Duplication	B	N	Severe	Not reported	151.993	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Pro1400Serfs*38	Pro1381Serfs*38	Frameshift	Duplication	B	N	Severe	Yes	122.752	N		Vencesla A, Corral-Rodriguez MA, Baena M, et al, Blood 111:3468-78,
p.(Pro1416Hisfs*8)	Pro1397Hisfs*8	Frameshift	Deletion	B	N	Severe	Not reported	64.295	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Asp1430Leufs*10)	Asp1411Leufs*10	Frameshift	Deletion	B	N	Severe	Not reported	57.217	N		David D, Ventura C, Moreira I, et al, Haematologica 91:840-3, 16769589
p.(Ser1445Leufs*20)	Ser1426Leufs*20	Frameshift	Deletion	B	N	Severe	Not reported	56.387	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Gly1456Glu)	Gly1437Glu	Missense	Substitution	B	N	Moderate	Not reported	23.488	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Lys1459Asnfs*6)	Lys1440Asnfs*6	Frameshift	Deletion/Insertion	B	Y	Severe	Not reported	194.523	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Asn1460Ilefs*5	Asn1441Ilefs*5	Frameshift	Deletion	B	Y	Severe	Yes	118.079	N		Tuddenham EG, Cooper DN, Gitschier J, et al, Nucleic Acids Research 19:4821-33, 1923751
p.Asn1460Lysfs*2	Asn1441Lysfs*2	Frameshift	Duplication	B	Y	Moderate/Severe	Yes	118.079	N		0, 0, 1908096
p.Asn1461Thrfs*13	Asn1442Thrfs*13	Frameshift	Deletion	B	Y	Moderate	Yes	124.064	N	Glycosylation site	Hua BL, Yan ZY, Liang Y, et al, Chinese Medical Journal 123:305-10, 20193250
p.Thr1468Profs*4	Thr1449Profs*4	Frameshift	Deletion	B	Y	Severe	Yes	10.517	N		Chen YC, Hu SH, Cheng SN, et al, Haemophilia 16:538-44, 20236351
p.Glu1470Valfs*94	Glu1451Valfs*94	Frameshift	Deletion	B	N	Severe	Yes	77.315	Y		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.(Gly1473Valfs*94)	Gly1454Valfs*94	Frameshift	Deletion	B	N	Severe	Yes	1.689	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Thr1483Glnfs*84	Thr1464Glnfs*84	Frameshift	Deletion	B	N	Severe	Yes	0.832	N		Reitter S, Sturn R, Horvath B, et al, Thrombosis & Haemostasis 104:78-85, 20431853
p.Ser1484Metfs*83	Ser1465Metfs*83	Frameshift	Deletion	B	N	Severe	Yes	0.08	N		Miller CH, Benson J, Ellingsen D, et al, Haemophilia 18:375-82, 22103590
p.Tyr1491*	Tyr1472*	Nonsense	Substitution	B	N	Severe	Yes	102.847	N		Green PM, Bagnall RD, Waseem NH, et al, British Journal of Haematology 143:115-28, 18691168
p.Lys1492*	Lys1473*	Nonsense	Substitution	B	N	Severe	Yes	114.335	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Pro1515Glnfs*52	Pro1496Glnfs*52	Frameshift	Deletion	B	N	Severe	Yes	35.858	N		Casana P, Cabrera N, Cid AR, et al, Haematologica 93:1091-4, 18403393
p.(Gly1547Glufs*20)	Gly1528Glufs*20	Frameshift	Deletion	B	N	Severe	Not reported	41.509	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Glu1549Glyfs*4)	Glu1530Glyfs*4	Frameshift	Deletion	B	N	Severe	Not reported	71.847	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Trp1554*	Trp1535*	Nonsense	Substitution	B	N	Severe	Yes	58.706	Y		Konkle BA, Fletcher SN, Johnsen JM, et al, My Life Our Future, unpublished,
p.Asn1555Lysfs*31	Asn1536Lysfs*31	Frameshift	Deletion/Insertion	B	N	Severe	Yes	2.707	N		Bogdanova N, Markoff A, Pollmann H, et al, Bogdanova N, Markoff A, Pollmann H, et al, 12204009
p.Asn1555Lysfs*13	Asn1536Lysfs*13	Frameshift	Deletion/Insertion	B	N	Severe	Yes	2.707	N		Bogdanova N, Markoff A, Pollmann H, et al, Bogdanova N, Markoff A, Pollmann H, et al, 12204009
p.(Phe1565*)	Phe1546*	Nonsense	Deletion	B	N	Severe	Not reported	200.994	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.(Lys1575Argfs*46)	Lys1556Argfs*46	Frameshift	Deletion	B	N	Severe	Not reported	150.512	N		CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. Human Mutation 34, E2382–E2392 (2013)
p.Trp1586*	Trp1567*	Nonsense	Substitution	B	N	Severe	Yes	78.797	N		Vidal F, Farssac E, Altisent C, et al, Vidal F, Farssac E, Altisent C, et al, 11341489
p.Trp1586*	Trp1567*	Nonsense	Substitution	B	N	Not reported	Yes	78.797	N		Vidal F, Farssac E, Altisent C, et al, Vidal F, Farssac E, Altisent C, et al, 11341489
p.Trp1599*	Trp1580*	Nonsense	Substitution	B	N	Severe	Yes	81.007	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Trp1599*	Trp1580*	Nonsense	Substitution	B	N	Severe	Yes	81.007	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Thr1609Asnfs*4	Thr1590Asnfs*4	Frameshift	Duplication	B	Y	Moderate/Severe	Yes	73.943	N		HAMSTERS (The Haemophilia A Mutation, Structure, Test and Resource Site), accessed 3/1/2011
p.Asp1615Glyfs*10	Asp1596Glyfs*10	Frameshift	Duplication	B	Y	Severe	Yes	147.179	N		Miller CH, Benson J, Ellingsen D, et al, HIRS, unpublished, CHAMP
p.Asn1633Lysfs*9	Asn1614Lysfs*9	Frameshift	Deletion	B	N	Severe	Yes	42.084	N		Vinciguerra C, Zawadzki C, Dargaud Y, et al, Thrombosis & Haemostasis 95:593-9, 16601827
p.Glu1634Glyfs*8	Glu1615Glyfs*8	Frameshift	Deletion	B	N	Severe	Yes	147.885	N		Miller CH, Benson J, Ellingsen D, et al, Haemophilia 18:375-82, 22103590
p.(Glu1642*)	Glu1623*	Nonsense	Substitution	B	N	Severe	Not reported	132.825	Y		Margaglione M, Castaman G, Morfini M, et al, Haematologica 93:722-8, 18387975
p.Trp1645*	Trp1626*	Nonsense	Substitution	B	N	Severe	Yes	184.232	N		Vinciguerra C, Zawadzki C, Dargaud Y, et al, Thrombosis & Haemostasis 95:593-9, 16601827
p.Trp1645*	Trp1626*	Nonsense	Substitution	B	N	Severe	Yes	184.232	N		Vinciguerra C, Zawadzki C, Dargaud Y, et al, Thrombosis & Haemostasis 95:593-9, 16601827
p.Ala1646Glnfs*17	Ala1627Glnfs*17	Frameshift	Deletion	B	N	Not reported	Yes	22.766	N		Rosset C, Gorziza RP, Botton MR, et al, Blood Coagulation & Fibrinolysis 25:125-7, 23963097

Table 14. B-domain mutations. Comparison of surface accessibility between *Gly/FC-FL-FVIII* and *Gly/BDD-FVIII* for residues involved in B-domain mutations causing severe/moderate Hemophilia-A and inhibitor formations mutations based on accessibility changes that affect binding sites and overall protein function

CHAPTER 4 - DISCUSSION

4.1 The FVIII B-domain wraps the central core of the FVIII protein.

Our AFM analysis of the rFL-FVIII, and rBDD-FVIII processed in both air and liquid revealed that each isolated surface to be globular, while pdFL-FVIII revealed globular shapes with tail-like extensions. [23,26,44,86,161] Specifically, rFL-FVIII globular structures disclosed high dense region in the center surrounding a light-scattered region extended in different directions. Measurements of these globular rFL-FVIII surfaces displayed approximately 6-7 nm height, and 30-40 nm width, while rBDD-FVIII showed approximately 4-5 nm height and 20-30 nm width, and the pdFL-FVIII had a globular height of 6-7 nm, a tail height of 3-4 nm, with a width of ~ 60-80 nm, suggesting that the B-domain contributes significantly to FVIII's architecture (Figure 31). Previous AFM studies have revealed that the rFL-FVIII molecules are generally spherical and a detailed view into it shows that some of these molecules are composed of core globular units with B-domain as serpentine tail structures and contain a series of small knots. In these, the presence of the B-domain was also found to be seen as isolated either globular or unraveled with a texture very similar to the tails observed previously which was explained as "peripheral satellite appendages", but our AFM imaging of rFL-FVIII suggests a wrapping of the B-domain without forming an independent appendage or domain. This wrapping structure hinted towards the shielding role of the B-domain. Furthermore, earlier studies also showed the effect of calcium concentration influencing the FL-FVIII structure displaying a calcium-dependent transition from a coiled to an unraveled tail structure of the FVIII B-domain, while this transition was not observed in our PBS-prepared FVIII. Certain AFM dockings showed the C1- and C2-domains oriented toward the negatively charged mica surface, mimicking FVIII's interaction with phosphatidylserine membranes (Figure 34). In the FL-FVIII-vWF complex topographs, globular portions resembled rFL-FVIII and tail-like structures likely represent clustered vWF multimers (Figure 31). Multimer analysis of all three purified peaks of pdFL-FVIII displayed the peak-2 corresponding to the FVIII-vWF complex to consist of dimeric, tetrameric, and hexameric complexes which meant that we had captured lower-middle molecular-weight multimers of vWF bound to FVIII in our AFM images (Figure 29). Our attempts to resolve the structure through cryo-EM also provided us the information on structural and conformational orientations of the B-domain and FL-FVIII in both recombinant and plasma-derived proteins. 62k particles of rFL-FVIII protein observed after the particle picking and extraction showed a high level of both structural and conformational heterogeneity. However, approximately 40k particles of the pdFL-FVIII protein showed a better structural homogeneity yet the conformational heterogeneity was observed at high levels (Figure 31). This confirms that the highly disordered B-domain contributes to the heterogeneity observed in these particles. Early transmission electron microscopy (TEM) studies displayed a short visible 14 nm tail in a negatively stained porcine FVIII specimen, which was assigned to the B-domain and later confirmed the presence of long tails using positive staining from human plasma-derived FVIII also suggesting a double-stranded rodlike shape of the B-domain. [48-50,162,163] It was also observed that the A- and C-domains of FVIII constitute a globular head, and the connecting B-domain is in a two-stranded tail. However, in our negative staining and cryo-EM studies, rFL-FVIII displayed structural and conformational diversity while pdFL-FVIII exhibited better structural uniformity, albeit with persistent conformational heterogeneity, likely due to the disordered B-domain (Figure 29 and Figure 31) Rigid body fitting revealed the B-domain appeared dispersed around the core of the FVIII protein (Figure 30). With these observations, we conclude that the FL-FVIII is a globular structure where the B-domain wraps around the core domains, only exposing the C-domains that bind to membranes and vWF, reinforcing its shielding role. Owing to its high disorderness and low resolution of the density maps, the structure of the B-domain and FL-FVIII protein remains unresolved but further attempts to increase the resolution are underway.

4.2 Critical roles of glycosylations, furin cleavage in the B-domain in maintaining the structure of FL-FVIII

The B-domain, rich in complex-type and high-mannose N-linked oligosaccharides, plays essential roles in FVIII intracellular processing and secretion. Key glycosylation sites (N1829, N2137, N2150) have been shown to facilitate interactions with ER chaperones Calnexin and Calreticulin and sorting receptors LMAN1 and MCFD2, supporting proper assembly of FVIII's structural domains. BDD-FVIII shows reduced binding to the LMAN1-MCFD2 complex, indicating that the B-domain plays an important role in cargo binding. [35,164] Additionally, these glycan structures are also implicated in the clearance pathways, where LRP and ASGPR bind to the B-domain with high affinity (Kd ≈ 2 nm). [41-43,165] Here we show the role of glycosylation concerning the structural and conformational stability of FL-FVIII. In our *Gly/FC-FL-FVIII* models, we observed that glycosylations significantly contribute to the structural and conformational stability of FVIII (minimal fluctuations in RMSF, RMSD, and Rg), corroborating earlier theories that carbohydrate moieties play a vital role in stabilizing folded domains (Figure 19). Glycosylation prevents aggregation, enhances stability through polar surfaces, and increases H-bonds, particularly in the B-domain. Glycosylation raises the free energy of the unfolded state, promoting stable conformations, facilitating specific interactions with carbohydrate-recognizing partners, and protecting FVIII from premature degradation. [166-168] Additionally, *FL-FVIII-vWF-D'-D3* complex model conformations with both primary (R1667) and secondary (R1332) furin cleaved indicates that the B-domain surrounding the central core of FVIII protein unravels as it passes through Golgi through the furin cleavages, thereby facilitating the binding of crucial interacting partners such as vWF (Figure 21, Figure 22). Additionally, our data supports the stabilizing effect of glycosylations on the FVIII protein aligns with the design of extended half-life (EHL) FVIII products like Rurioctocog alfa pegol (Adynovate). In these EHL products, the B-domain and glycosylation sites are strategically modified, often through PEGylation, to prolong FVIII's plasma half-life. By attaching PEG moieties to the B-domain, receptor-mediated clearance is minimized, which stabilizes FVIII in the bloodstream, allowing for extended dosing intervals and enhanced therapeutic outcomes in the treatment of HA. [169] Our analysis of the AF-FL-FVIII model without furin cleavage reveals notable structural tension and instability, indicating that furin cleavage may be essential for alleviating this tension and promoting stability in the full-length FVIII structure. Prior studies have shown that glycosylation and furin cleavage together greatly enhance protein stability and activity, facilitating a stable conformation evident in our *Gly/FC-FL-FVIII* model, highlighting its reliability. [115] Moreover, the furin cleavage further promotes the unwrapping of the B-domain around FVIII during Golgi processing, facilitating interactions with binding partners like vWF.

4.3 Shielding effect of the B-domain

As explained earlier in Section 4.1, we strongly support that the B-domain wrapping around the central core of the FVIII protein offers a shielding effect. This shielding effect is important for stabilizing FVIII and preventing premature activation. It is believed that the presence of the B-domain could slow down the kinetics of activation of full-length FVIII compared to BDD-FVIII and we strongly support the idea that the B-domain plays an important role in regulating the activation of FVIII. Previous structural studies on C1- and C2-domains have elucidated that these regions are rich in several basic and hydrophobic residues that are critical for their ability to interact with phospholipid membrane surfaces. In addition to this, the C-domains have also been shown to have a high isoelectric point (9.2) and are positively charged allowing the binding of FVIII to the membrane surfaces through electrostatic and hydrophobic interactions. To further elucidate this, differences in the continuum electrostatics studies in the presence and absence of glycosylated B-domain displayed significant differences between the same (Figure 34). Our refined *Gly/FC-FL-FVIII* model differs from the source AF-FL-FVIII model since in the source model B-domain hinders access to key functional regions in the C-domains. However, the *Gly/FC-FL-FVIII* model maintains the pseudo-three-fold symmetry in the A-domains and allows for unobstructed access to the functionally important regions within the C-domains. Additionally, we infer the shielding mechanism of the B-domain based on the steric hindrance and physical overlap. The positively charged C1- and C2-domains were observed to retain their charge in both *Gly/FC-FL-FVIII* and *Gly-BDD-FVIII* suggesting that the presence of the B-domain enables direct interactions with critical binding partners such as vWF, thrombin, and PS surfaces, preserving FVIII's functional integrity (Figure 34). [23,151]

a. **Inhibitors:** The development of an anti-FVIII immune response is a major complication associated with the treatment of severe hemophilia A patients. Being one of the highly immunogenic proteins, significant efforts have been made to develop FVIII constructs that maintain longer half-lives, increased activity, and decreased immunogenicity. Currently, approximately 30% of severe hemophilia A patients receiving FVIII replacement therapy develop this complication and A2-, C1-, and C2-domains harbor major immunogenic responses that are recognized by the inhibitory antibodies. Based on the mechanism of FVIII inhibition the antibody inhibitors developed are classified into the “classical” antibody inhibitors that block the binding of FVIII to vWF or phospholipid surfaces and “non-classical” antibody inhibitors that block the proteolytic activation of FVIII by thrombin or FXa, thus preventing the dissociation of FVIII from vWF. The functional characteristics of the A2-domain include the binding to the FXa and the C1- and C2-domains contain the interaction sites necessary for binding to the vWF and phospholipid membrane surfaces. Comparative SASA studies reveal the B-domain’s shielding role for inhibitor binding i.e. by limiting epitope accessibility. For anti-C2 inhibitors, B-domain residues overlap Ab binding sites, potentially neutralizing their impact on FL-FVIII. Similarly, anti-C1 Abs (ET3i/2A9, NB2E9, 8TY1, NB33, NBG12) and anti-A3, anti-A2 inhibitors show reduced membrane-site accessibility in the B-domain’s presence. Previous research on inhibitors has demonstrated that the R503-I527 region interacts with FIXa and LRP is highly immunogenic. The decreased accessibility of FVIII in the presence of the B-domain suggests a protective effect, neutralizing and reducing the impact of inhibitor binding on FVIII activation (Figure 35 and Table 11). [30,36,69-72,72,73,148,149]

b. **Shielding from interacting partners:** In our FL-FVIII model, the B-domain partially covers the A2 domain, blocking access to thrombin/FXa cleavage sites. We propose that the B-domain provides stabilization by wrapping around these labile regions. SASA variations between the *Gly/FC-FL-FVIII* model and the BDD-FVIII structure suggest that the B-domain shields thrombin cleavage sites (R391, R759, R1708), and thrombin-binding regions in A1- and A3-domains. The first cleavage site (R759), with its notably reduced accessibility compared to the other two sites, might be instrumental in preventing premature activation thus affecting FVIII activation kinetics (Figure 36, Table 12). [24,29,36,38,132] Concerning vWF, as FVIII in plasma is present as a complex that is bound non-covalently to vWF, this complex stabilizes the FVIII, and protects FVIII from premature clearance and activation, ensuring its survival in circulation. We strongly support that furin cleavage is pivotal for B-domain release, enabling vWF binding. [33] Previous studies showed the primary furin cleavage unlocks the region where vWF-D’-D3 is known to bind and *FL-FVIII-vWF-D’-D3* complex model showed that the primary furin cleavage (R1667) exposes binding regions by removing certain B-domain residues that sterically interfere with vWF-D’ binding. The secondary cleavage (R1332) partially unwraps the B-domain allowing efficient and stable vWF binding (Figure 22, Figure 35) (Table 12). [44,47,114,116-118,132]

c. **Mutations:** Several mutations impacting thrombin cleavage, vWF binding, and tyrosine sulfation were observed to be buried within the B-domain, limiting antibody binding to affected epitopes (Figure 37, Figure 38, Table 13, and Table 14). Notably, missense mutations in the B-domain, such as p.G1456E, may destabilize the protein due to glycine's unique role in maintaining flexibility and proper folding. The p.F1194V mutation, which affects interactions with the A1-domain signal peptide, likely impairs FVIII secretion, while p.L1182V at the interface may weaken overall structural stability (Table 14). Mutations on or near glycosylation sites (p.M798T mutation) that impair glycosylation may further hasten FVIII degradation. Missense B-domain mutations located at its interface with the A1-domain, appear to disrupt critical interactions that maintain FVIII's structural integrity. For instance, the mutations p.L1182V and p.F1194V were identified at the B-domain-A1-domain interface, where they disrupt hydrogen bonding and introduce steric clashes that could destabilize the domain interface (Figure 38, Table 14). This disruption likely impairs proper alignment and structural cohesion between the B- and A1-domains, potentially compromising FVIII function. Similarly, the p.G1437Q mutation introduces steric clashes with surrounding residues in the B-domain, which could hinder the local structural flexibility necessary for FVIII's stability. Our FL-FVIII model shows an overall reduction in SASA for inhibitor epitopes, and various mutations (irrespective of the type of mutational pathogenicity) in the presence of the B-domain, causing inhibitory Abs to bind less efficiently to FL-FVIII. However, this effect can vary and is subjective. In some cases, the mutated epitope, together with the B-domain, may form a neo-epitope that could provoke a stronger inhibitory response. [51,53,56,57]

CONCLUSION

In conclusion, our study emphasizes the critical role of the FVIII B-domain in maintaining the structure and function of FL-FVIII. The B-domain shields important domains, reducing accessibility to inhibitors, interacting proteins, and cleavage sites, thereby stabilizing FVIII. Glycosylation on the B-domain enhances FVIII's structural stability by preventing aggregation and promoting proper folding. Through this work, we have taken a critical step forward in understanding and elucidating the structural-functional role of the B-domain within FL- FVIII. Future work should explore B-domain modifications that further enhance immune tolerance and stability, aiding in the development of next-generation therapies for Hemophilia A. This research underscores the potential of AI and AlphaFold2 to expedite the study of structural diversity, facilitating the development of novel biologics and advancing protein and drug design.

FUTURE PERSPECTIVES

The future direction of this project could involve expanding cryo-electron microscopy trials to achieve atomic resolution of the full-length FVIII structure. However, due to the inherent flexibility and disorder of the B-domain, fully resolving it may be challenging. Additionally, cross-linking mass spectrometry could be utilized to gain detailed intra- and inter-domain structural insights. Given the heavy glycosylation of the B-domain, optimizing deglycosylation protocols without compromising protein integrity will be essential for effective cross-linking. Moreover, structural characterization can be extended to study FVIII inhibitors and interacting partners such as LMAN1, vWF, and thrombin. These studies could stabilize the FVIII-inhibitor complex and provide valuable insights into the unknown regulatory pathways of full-length FVIII.

ACKNOWLEDGEMENTS

I would like to extend my heartfelt gratitude to **Prof. Dr. Johannes Oldenburg** for allowing me to work at the Institute of Experimental Haematology and Transfusion Medicine. Obliged to second supervisor, **Prof. Dr. Diana Imhof**, for her co-supervision.

Thankful to my primary supervisor, **PD Dr. Arijit Biswas**, for entrusting me with this project. This project was funded by the IIR-DE-002689-BISWAS awarded to PD Dr. Arijit Biswas, and Prof. Dr. Johannes Oldenburg from Takeda (Shire). I also thank the lab postdoc—**Dr. Sneha Singh.**

Incredibly grateful to the people who started as colleagues but are now the truest confidants, **Mr. Haroon Javed and Mr. Mubasshurel Islam,** who stood by my side through some hard times and joyful moments.

A special note of gratitude to **Dr. Jean-Luc Pellequer**, who has been an exceptional supervisor. His unwavering support/guidance in building and refining the structural models has been a core part of this project. Deeply grateful to him for introducing me to the fascinating world of AFM and rekindling my passion for structural biology. Many thanks to **Dr. and Mrs. Claire**, for being so welcoming and making me feel homely in Grenoble. The expenses for the AFM work conducted at the Institut de Biologie Structurale in Grenoble, France were funded and supported by the Society for Thrombosis and Haemostasis Research e.V (GTH Rudolf-Marx-Stipendium) to Ms. Samhitha Urs, Ramaraje Urs.

Thankful to **Prof. Jens Müller** and **Ms. Simone Gaesper**, for helping with the FVIII multimer analysis. It has been a truly valuable detail to the project. Grateful to **Ms. Deniz Ugurlar and Ms. Boxue Ma** for performing a detailed cryoEM analysis of the FVIII protein at ThermoFischer Netherlands.

I also wish to express my sincere thanks to my colleagues, **PD Dr. Behnaz Pezeshkpoor** and **PD Dr. Hamideh Yadegari**, for recognizing my potential in protein biology and providing the opportunities to contribute to a small part of their work leading to two successful publications.

To the people close to my heart, **Ms. Rawya Al-Rifai** and **Ms. Mira Ibrahim,** thank you for being more than friends—you have truly been my sisters and my family here in Germany. You both have made my PhD days not just bearable, but meaningful!

No words will ever be enough to express how thankful I am to my family - My grandparents, Parents, and my brother. My partner - my constant, my source of calm in these four years of chaos, and my greatest cheerleader— Dr. Abhishek, words fall behind how indebted I am to you! ***All that I am, all I hope to be, and all I will ever be in this life, I owe it to them.***

Infinite gratitude and salutations to the Almighty.

CONFERENCE AND WORKSHOP PARTICIPATIONS

ORAL PRESENTATIONS:

- **International Society on Thrombosis and Haemostasis, Inc - ISTH 2022 (London)**
  Title: Exploring the conformational landscape of Factor VIII B domain in order to generate an all atom full length structure of the coagulation Factor VIII protein.
- **55. Jahrestagung der Deutschen Gesellschaft für Transfusionsmedizin und Immunhämatologie e.V. (DGTI) - 2022 (Mannheim)**
  Title: Structural investigations into coagulation factor VIII full-length and B domain
- **67. Jahrestagung der GTH Gesellschaft für Thrombose- und Hämostaseforschung e.V. - 2023 (Frankfurt)**
  Title: Low resolution cryo-EM maps and AFM analysis combined with alpha fold model of full-length coagulation Factor VIII sheds light on the conformational positioning of the Factor VIII B domain.
- **International Society on Thrombosis and Haemostasis, Inc - ISTH 2023 (Montreal)**
  Title: Structural characterization of coagulation factor VIII

POSTER PRESENTATIONS:

- **52. Hamburger Hämophilie Symposion - HHS 2021 - Virtual**
  Title: Structural Characterization of Factor VIII B Domain to Generate an All Atom Full-Length Structure of the Coagulation Factor VIII Protein.
  Ramaraje Urs , Samhitha Urs[1] , Sneha Singh[1] , Haroon Javed[1] , Guy Schoehn[2] , Jean-Luc Pellequer[2] , Jean-Marie Teulon[2] , Daphne Fenel[2] , Johannes Oldenburg[1] , Arijit Biswas[1] - Hamostaseologie 2022; 42(S 01): S32-S37, DOI: 10.1055/s-0042-1758498
- **66. Jahrestagung der GTH Gesellschaft für Thrombose- und Hämostaseforschung e.V. - 2022 (Leipzig)**
  Title: Structural investigations into coagulation factor VIII B domain
- **53. Hamburger Hämophilie Symposion - HHS 2022 - Hamburg**
  Title: Low resolution cryoEM combined with alpha fold model of full length coagulation Factor FVIII sheds light on the spatial orientation of B domain.
- **68. Jahrestagung der GTH Gesellschaft für Thrombose- und Hämostaseforschung e.V. - 2024 (Wein)**
  Title: In-detail AFM analysis of the different conformational states of full-length coagulation Factor VIII and FVIII-vWF complex.
  S. U. Ramaraje Urs[1], J. L. Pellequer[2], J. M. Teulon[2] , S. Singh[1] , J. Oldenburg[1] , A. Biswas[1] - Hamostaseologie 2024; 44(S 01): S44-S45 DOI: 10.1055/s-0044-1779121
- **International Society on Thrombosis and Haemostasis, Inc - ISTH 2024 (Bangkok)**
  Title: Understanding distinct conformational states of full-length coagulation Factor VIII and its complex with vWF: A comprehensive AFM analysis

SUMMER SCHOOL AND WORKSHOPS:

- **Writing Grant Proposals** - 08 July 2021 - Bonn Graduate Center (BGZ), University of Bonn
- **Mass Spectrometry in Life Sciences -** 29 September-01 October, BIGS DrugS, University of Bonn
- **Beyond the PhD - What's Next? How to Find and Reach Your Goals** - 01-03 February 2022 - Bonn Graduate Center (BGZ), University of Bonn
- **Hands-on Introduction to R** - 10-12 October 2022 - Core Unit for Bioinformatics Analysis (CUBA), University of Bonn
- **AFM BioMed Summer School 2023** - 21-26 August 2023, Marcoule, France
- **Introduction to Programming with Python** - 25-26 April 2024, High Performance Computing Team at HRZ, University of Bonn

INVITED TALKS:

- Invited internal scientific webinar on **"Structure of full-length FVIII with focus on the B-domain"** - on 19.06.2023 in and by **Takeda Pharmaceutical Company Limited** (across U.S., Germany, and Austria)- Peter L. Turecek, Hon. Prof.(FH) Univ.-Doz.

ACHIEVEMENTS:

- **Reisestipendium** - 55. Jahrestagung der Deutschen Gesellschaft für Transfusionsmedizin und Immunhämatologie e.V. (DGTI) - 2022 (Mannheim)
- **Poster award** - 53. Hamburger Hämophilie Symposion - HHS 2022 - Hamburg
- **Rudolf Marx Stipendium** - 67. Jahrestagung der GTH Gesellschaft für Thrombose- und Hämostaseforschung e.V. - 2023 (Frankfurt)
- **Early Career Travel Award** - International Society on Thrombosis and Haemostasis, Inc - ISTH 2023 (Montreal)

Student Representative: Bonn International Graduate School of Drug Sciences (BIGS DrugS) - 25.06.2024 - Present

FIGURES:

TABLES:

REFERENCES:

1. Butenas, Mann, & Butenas. [No title found]. *Biochemistry (Moscow)* **67**, 3-12 (2002).
2. Tanaka, K. A., Key, N. S. & Levy, J. H. Blood Coagulation: Hemostasis and Thrombin Regulation. *Anesthesia & Analgesia* **108**, 1433-1446 (2009).
3. Hoffman, M. & Monroe, D. M. Coagulation 2006: A Modern View of Hemostasis. *Hematology/Oncology Clinics of North America* **21**, 1-11 (2007).
4. Alberio & Dale. Platelet-collagen interactions: membrane receptors and intracellular signalling pathways. *Eur J Clin Investigation* **29**, 1066-1076 (1999).
5. Briedé, J., Heemskerk, J., Cornelis Van'T, V., Hemker, H. C. & Lindhout, T. Contribution of Platelet-derived Factor Va to Thrombin Generation on Immobilized Collagen- and Fibrinogen-adherent Platelets. *Thromb Haemost* **85**, 509-513 (2001).
6. Macfarlane, R. G. An Enzyme Cascade in the Blood Clotting Mechanism, and its Function as a Biochemical Amplifier. *Nature* **202**, 498-499 (1964).
7. Lorand, L. Factor XIII: Structure, Activation, and Interactions with Fibrinogen and Fibrin. *Annals of the New York Academy of Sciences* **936**, 291-311 (2001).
8. Nurden, A. Platelets, inflammation and tissue regeneration. *Thromb Haemost* **105**, S13-S33 (2011).
9. Dahlbäck, B. Progress in the Understanding of the Protein C Anticoagulant Pathway. *International Journal of Hematology* **79**, 109-116 (2004).
10. Bevers, E. M., Comfurius, P., Hemker, H. C. & Zwaal, R. F. A. On the procoagulant activity of platelets stimulated by collagen and thrombin. *Thrombosis Research* **33**, 553-554 (1984).
11. Gitschier, J. *et al.* Characterization of the human factor VIII gene. *Nature* **312**, 326-330 (1984).
12. Figueiredo, M. S. & Brownlee, G. G. cis-Acting Elements and Transcription Factors Involved in the Promoter Activity of the Human Factor VIII Gene. *Journal of Biological Chemistry* **270**, 11828-11838 (1995).
13. Fay, P. Factor VIII Structure and Function. *International Journal of Hematology* **83**, 103-108 (2006).
14. Coppola, A. *et al.* Factor VIII gene (F8) mutations as predictors of outcome in immune tolerance induction of hemophilia A patients with high-responding inhibitors. *Journal of Thrombosis and Haemostasis* **7**, 1809-1815 (2009).
15. Thompson, A. R. Structure and Function of the Factor VIII Gene and Protein. *Semin Thromb Hemost* **29**, 011-022 (2003).
16. Marquette, K. A., Pittman, D. D. & Kaufman, R. J. A 110-amino Acid Region within the A1-domain of Coagulation Factor VIII Inhibits Secretion from Mammalian Cells. *Journal of Biological Chemistry* **270**, 10297-10303 (1995).
17. Lapan, K. & Fay, P. Interaction of the A1 Subunit of Factor VIIIa and the Serine Protease Domain of Factor X Identified by Zero-length Cross-linking. *Thromb Haemost* **80**, 418-422 (1998).
18. Miao, H. Z. *et al.* Bioengineering of coagulation factor VIII for improved secretion. *Blood* **103**, 3412-3419 (2004).
19. Becker, S. *et al.* Confocal microscopy analysis of native, full length and B-domain deleted coagulation factor VIII trafficking in mammalian cells. *Thromb Haemost* **92**, 23-35 (2004).
20. Bloem, E. *et al.* Factor VIII with a 237 amino acid B-domain has an extended half-life in F8-knockout mice. *Journal of Thrombosis and Haemostasis* **17**, 350-360 (2019).
21. Anzengruber, J. *et al.* How Full-Length FVIII Benefits from Its Heterogeneity - Insights into the Role of the B-Domain. *Pharm Res* **36**, 77 (2019).
22. Grushin, K. *et al.* Lack of recombinant factor VIII B-domain induces phospholipid vesicle aggregation: implications for the immunogenicity of factor VIII. *Haemophilia* **20**, 723-731 (2014).
23. Stoilova-McPhie, S., Lynch, G. C., Ludtke, S. & Pettitt, B. M. Domain organization of membrane-bound factor VIII: Membrane-Bound FVIII. *Biopolymers* **99**, 448-459 (2013).
24. Pahl, S., Pavlova, A., Driesen, J. & Oldenburg, J. Effect of F8 B domain gene variants on synthesis, secretion, activity and stability of factor VIII protein. *Thromb Haemost* **111**, 58-66 (2014).
25. Gilbert, G. E., Kaufman, R. J., Arena, A. A., Miao, H. & Pipe, S. W. Four Hydrophobic Amino Acids of the Factor VIII C2 Domain Are Constituents of Both the Membrane-binding and von Willebrand Factor-binding Motifs. *Journal of Biological Chemistry* **277**, 6374-6381 (2002).
26. Shen, B. W. *et al.* The tertiary structure and domain organization of coagulation factor VIII. *Blood* **111**, 1240-1247 (2008).
27. Spiegel, P. C., Jacquemin, M., Saint-Remy, J.-M. R., Stoddard, B. L. & Pratt, K. P. Structure of a factor VIII C2 domain-immunoglobulin G4κ Fab complex: identification of an inhibitory antibody epitope on the surface of factor VIII. *Blood* **98**, 13-19 (2001).
28. Pratt, K. P. *et al.* Structure of the C2 domain of human factor VIII at 1.5 Å resolution. *Nature* **402**, 439-442 (1999).

29. Nogami, K. *et al.* Factor VIII C2 Domain Contains the Thrombin-binding Site Responsible for Thrombin-catalyzed Cleavage at Arg1689. *Journal of Biological Chemistry* **275**, 25774-25780 (2000).
30. Smith, I. W. *et al.* The 3.2 Å structure of a bioengineered variant of blood coagulation factor VIII indicates two conformations of the C2 domain. *Journal of Thrombosis and Haemostasis* **18**, 57-69 (2020).
31. Ngo, J. C. K., Huang, M., Roth, D. A., Furie, B. C. & Furie, B. Crystal Structure of Human Factor VIII: Implications for the Formation of the Factor IXa-Factor VIIIa Complex. *Structure* **16**, 597-606 (2008).
32. Svensson, L. A., Thim, L., Olsen, O. H. & Nicolaisen, E. M. Evaluation of the metal binding sites in a recombinant coagulation factor VIII identifies two sites with unique metal binding properties. *Biological Chemistry* **394**, 761-765 (2013).
33. Fuller, J. R., Knockenhauer, K. E., Leksa, N. C., Peters, R. T. & Batchelor, J. D. Molecular determinants of the factor VIII/von Willebrand factor complex revealed by BIVV001 cryo-electron microscopy. *Blood* **137**, 2970-2980 (2021).
34. Leksa, N. C. *et al.* The structural basis for the functional comparability of factor VIII and the long-acting variant recombinant factor VIII Fc fusion protein. *Journal of Thrombosis and Haemostasis* **15**, 1167-1179 (2017).
35. Pipe, S. W. Functional roles of the factor VIII B domain. *Haemophilia* **15**, 1187-1196 (2009).
36. Arsiccio, A., Beavis, J., Raut, S. & Coxon, C. H. FVIII inhibitors display FV-neutralizing activity in the prothrombin time assay. *Journal of Thrombosis and Haemostasis* **19**, 1907-1913 (2021).
37. Hultin, M. Modulation of thrombin-mediated activation of factor VIII:C by calcium ions, phospholipid, and platelets. *Blood* **66**, 53-58 (1985).
38. Eaton, D., Rodriguez, H. & Vehar, G. A. Proteolytic processing of human factor VIII. Correlation of specific cleavages by thrombin, factor Xa, and activated protein C with activation and inactivation of factor VIII coagulant activity. *Biochemistry* **25**, 505-512 (1986).
39. Myles, T., Yun, T. H. & Leung, L. L. K. Structural requirements for the activation of human factor VIII by thrombin. *Blood* **100**, 2820-2826 (2002).
40. Khrenov, A. V., Ananyeva, N. M. & Saenko, E. L. Role of the B domain in proteolytic inactivation of activated coagulation factor VIII by activated protein C and activated factor X. *Blood Coagulation & Fibrinolysis* **17**, 379-388 (2006).
41. Bovenschen, N. *et al.* Low Density Lipoprotein Receptor-related Protein and Factor IXa Share Structural Requirements for Binding to the A3 Domain of Coagulation Factor VIII. *Journal of Biological Chemistry* **278**, 9370-9377 (2003).
42. Bovenschen, N., Van Stempvoort, G., Voorberg, J., Mertens, K. & Meijer, A. B. Proteolytic cleavage of factor VIII heavy chain is required to expose the binding-site for low-density lipoprotein receptor-related protein within the A2 domain. *Journal of Thrombosis and Haemostasis* **4**, 1487-1493 (2006).
43. Bovenschen, N., Rijken, D. C., Havekes, L. M., Van Vlijmen, B. J. M. & Mertens, K. The B domain of coagulation factor VIII interacts with the asialoglycoprotein receptor. *Journal of Thrombosis and Haemostasis* **3**, 1257-1265 (2005).
44. Pipe, S. W., Montgomery, R. R., Pratt, K. P., Lenting, P. J. & Lillicrap, D. Life in the shadow of a dominant partner: the FVIII-VWF association and its clinical implications for hemophilia A. *Blood* **128**, 2007-2016 (2016).
45. Clifton, J. G. *et al.* Proteomic characterization of plasma-derived clotting factor VIII-von Willebrand factor concentrates. *Electrophoresis* **30**, 3636-3646 (2009).
46. Cao, W., Cao, W., Zhang, W., Zheng, X. L. & Zhang, X. F. Factor VIII binding affects the mechanical unraveling of the A2 domain of von Willebrand factor. *Journal of Thrombosis and Haemostasis* **18**, 2169-2176 (2020).
47. Drago, V. *et al.* In-Silico Characterization of von Willebrand Factor Bound to FVIII. *Applied Sciences* **12**, 7855 (2022).
48. Hamer, R. J. *et al.* Factor VIII binds to von Willebrand factor via its Mr-80000 light chain. *Eur J Biochem* **166**, 37-43 (1987).
49. Lollar, P. & Parker, C. G. Stoichiometry of the porcine factor VIII-von Willebrand factor association. *Journal of Biological Chemistry* **262**, 17572-17576 (1987).
50. Sakariassen, K., Ottenhof-Rovers, M. & Sixma, J. Factor VIII-von Willebrand factor requires calcium for facilitation of platelet adherence. *Blood* **63**, 996-103 (1984).
51. Eckhardt, C. L. *et al.* Factor VIII gene (F8) mutation and risk of inhibitor development in nonsevere hemophilia A. *Blood* **122**, 1954-1962 (2013).
52. Wacey, A. The haemophilia A mutation search test and resource site, home page of the factor VIII mutation database: HAMSTeRS. *Nucleic Acids Research* **24**, 100-102 (1996).
53. Wei, W. *et al.* Missense mutations near the N-glycosylation site of the A2 domain lead to various intracellular trafficking defects in coagulation factor VIII. *Sci Rep* **7**, 45033 (2017).

54. Gouw, S. C. *et al.* F8 gene mutation type and inhibitor development in patients with severe hemophilia A: systematic review and meta-analysis. *Blood* **119**, 2922-2934 (2012).
55. Payne, A. B., Miller, C. H., Kelly, F. M., Michael Soucie, J. & Craig Hooper, W. The CDC Hemophilia A Mutation Project (CHAMP) Mutation List: A New Online Resource. *Human Mutation* **34**, E2382-E2392 (2013).
56. Jourdy, Y. *et al.* Characterization of five associations of *F8* missense mutations containing FVIII B domain mutations. *Haemophilia* **22**, 583-589 (2016).
57. Liu, M.-L. *et al.* Hemophilic factor VIII C1- and C2-domain missense mutations and their modeling to the 1.5-angstrom human C2-domain crystal structure. *Blood* **96**, 979-987 (2000).
58. Gouw, S. C. *et al.* Factor VIII Products and Inhibitor Development in Severe Hemophilia A. *N Engl J Med* **368**, 231-239 (2013).
59. Weyand, A. C. & Pipe, S. W. New therapies for hemophilia. *Blood* **133**, 389-398 (2019).
60. Marchesini, E., Morfini, M. & Valentino, L. Recent Advances in the Treatment of Hemophilia: A Review. *BTT* **Volume 15**, 221-235 (2021).
61. Calvez, T. *et al.* Recombinant factor VIII products and inhibitor development in previously untreated boys with severe hemophilia A. *Blood* **124**, 3398-3408 (2014).
62. Coppola, A. Treatment of hemophilia: a review of current advances and ongoing issues. *JBM* 183 (2010) doi:10.2147/JBM.S6885.
63. Lieuw, K. Many factor VIII products available in the treatment of hemophilia A: an embarrassment of riches? *JBM* **Volume 8**, 67-73 (2017).
64. Wu, J., Liu, X., Yang, H., He, Y. & Yu, D. Advances in biopharmaceutical products for hemophilia. *iScience* **27**, 111436 (2024).
65. Tomeo, F. *et al.* Haemophilia, state of the art and new therapeutic opportunities, a regulatory perspective. *Brit J Clinical Pharma* **87**, 4183-4196 (2021).
66. Lisowski, L., Staber, J. M., Wright, J. F. & Valentino, L. A. The intersection of vector biology, gene therapy, and hemophilia. *Research and Practice in Thrombosis and Haemostasis* **5**, e12586 (2021).
67. Nguyen, N. H., Jarvi, N. L. & Balu-Iyer, S. V. Immunogenicity of Therapeutic Biological Modalities - Lessons from Hemophilia A Therapies. *Journal of Pharmaceutical Sciences* **112**, 2347-2370 (2023).
68. Astermark, J. FVIII inhibitors: pathogenesis and avoidance. *Blood* **125**, 2045-2051 (2015).
69. Childers, K. C., Peters, S. C. & Spiegel, P. C. Structural insights into blood coagulation factor VIII: Procoagulant complexes, membrane binding, and antibody inhibition. *Journal of Thrombosis and Haemostasis* **20**, 1957-1970 (2022).
70. Childers, K. C. *et al.* Structure of coagulation factor VIII bound to a patient-derived anti-C1 domain antibody inhibitor. *Blood Journal* blood.2023020181 (2023) doi:10.1182/blood.2023020181.
71. Gish, J. S. *et al.* Structure of blood coagulation factor VIII in complex with an anti-C1 domain pathogenic antibody inhibitor. *Blood* **137**, 2981-2986 (2021).
72. Walter, J. D. *et al.* Structure of the factor VIII C2 domain in a ternary complex with 2 inhibitor antibodies reveals classical and nonclassical epitopes. *Blood* **122**, 4270-4278 (2013).
73. Spiegel, P. C., Jacquemin, M., Saint-Remy, J.-M. R., Stoddard, B. L. & Pratt, K. P. Structure of a factor VIII C2 domain-immunoglobulin G4κ Fab complex: identification of an inhibitory antibody epitope on the surface of factor VIII. *Blood* **98**, 13-19 (2001).
74. Burley, S. K. Integrative/Hybrid Methods Structural Biology: Role of Macromolecular Crystallography. in *Integrative Structural Biology with Hybrid Methods* (eds. Nakamura, H., Kleywegt, G., Burley, S. K. & Markley, J. L.) vol. 1105 11-18 (Springer Singapore, Singapore, 2018).
75. Meller, J. Molecular Dynamics. in *Encyclopedia of Life Sciences* (Wiley, 2001). doi:10.1038/npg.els.0003048.
76. Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **99**, 1129-1143 (2018).
77. Badar, M. S., Shamsi, S., Ahmed, J. & Alam, Md. A. Molecular Dynamics Simulations: Concept, Methods, and Applications. in *Transdisciplinarity* (ed. Rezaei, N.) vol. 5 131-151 (Springer International Publishing, Cham, 2022).
78. Emsley, P. & Cowtan, K. *Coot* : model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
79. Chen, V. B. *et al. MolProbity* : all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
80. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* **35**, W375-W383 (2007).
81. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* **27**, 293-315 (2018).

82. Fernandez-Leiro, R. & Scheres, S. H. W. Unravelling biological macromolecules with cryo-electron microscopy. *Nature* **537**, 339-346 (2016).
83. Benjin, X. & Ling, L. Developments, applications, and prospects of cryo-electron microscopy. *Protein Science* **29**, 872-882 (2020).
84. Sigworth, F. J. Principles of cryo-EM single-particle image processing. *Microscopy (Tokyo)* **65**, 57-67 (2016).
85. van Zundert, G. C. P., Melquiond, A. S. J. & Bonvin, A. M. J. J. Integrative Modeling of Biomolecular Complexes: HADDOCKing with Cryo-Electron Microscopy Data. *Structure* **23**, 949-960 (2015).
86. Bonazza, K. *et al.* Visualization of a protein-protein interaction at a single-molecule level by atomic force microscopy. *Anal Bioanal Chem* **406**, 1411-1421 (2014).
87. Chen, S. W., Teulon, J.-M., Godon, C. & Pellequer, J.-L. Atomic force microscope, molecular imaging, and analysis: Atomic Force Microscope, Molecular Imaging, and Analysis. *J. Mol. Recognit.* **29**, 51-55 (2016).
88. Pellequer, J. Perspectives Toward an Integrative Structural Biology Pipeline With Atomic Force Microscopy Topographic Images. *J of Molecular Recognition* e3102 (2024) doi:10.1002/jmr.3102.
89. Trinh, M.-H. *et al.* Computational Reconstruction of Multidomain Proteins Using Atomic Force Microscopy Data. *Structure* **20**, 113-120 (2012).
90. Amyot, R., Marchesi, A., Franz, C. M., Casuso, I. & Flechsig, H. Simulation atomic force microscopy for atomic reconstruction of biomolecular structures from resolution-limited experimental images. *PLoS Comput Biol* **18**, e1009970 (2022).
91. Chen, S. W., Banneville, A.-S., Teulon, J.-M., Timmins, J. & Pellequer, J.-L. Nanoscale surface structures of DNA bound to *Deinococcus radiodurans* HU unveiled by atomic force microscopy. *Nanoscale* **12**, 22628-22638 (2020).
92. Xia, F. & Youcef-Toumi, K. Review: Advanced Atomic Force Microscopy Modes for Biomedical Research. *Biosensors* **12**, 1116 (2022).
93. Jalili, N. & Laxminarayana, K. A review of atomic force microscopy imaging systems: application to molecular metrology and biological sciences. *Mechatronics* **14**, 907-945 (2004).
94. Vallat, B. *et al.* IHMCIF: An Extension of the PDBx/mmCIF Data Standard for Integrative Structure Determination Methods. *Journal of Molecular Biology* **436**, 168546 (2024).
95. Rout, M. P. & Sali, A. Principles for Integrative Structural Biology Studies. *Cell* **177**, 1384-1403 (2019).
96. Srivastava, A., Tiwari, S. P., Miyashita, O. & Tama, F. Integrative/Hybrid Modeling Approaches for Studying Biomolecules. *Journal of Molecular Biology* **432**, 2846-2860 (2020).
97. Sali, A. From integrative structural biology to cell biology. *Journal of Biological Chemistry* **296**, 100743 (2021).
98. Yu, L., Yang, D., Chu, M. & Sun, Y. Advances and challenges in the purification of recombinant coagulation factors: A review. *Journal of Chromatography A* **1716**, 464662 (2024).
99. Kumar, S. R. Industrial production of clotting factors: Challenges of expression, and choice of host cells. *Biotechnology Journal* **10**, 995-1004 (2015).
100. Blasi, B., D'Amici, G. M., D'Alessandro, A. & Zolla, L. Native analysis of plasma-derived clotting factor VIII concentrates: "Sponge effect" and contaminants. *Electrophoresis* **33**, 1292-1298 (2012).
101. D'Amici, G. M., Timperio, A. M., Gevi, F., Grazzini, G. & Zolla, L. Recombinant clotting factor VIII concentrates: Heterogeneity and high-purity evaluation. *Electrophoresis* **31**, 2730-2739 (2010).
102. Beveridge, R. & Calabrese, A. N. Structural Proteomics Methods to Interrogate the Conformations and Dynamics of Intrinsically Disordered Proteins. *Front. Chem.* **9**, 603639 (2021).
103. Eliezer, D. Biophysical characterization of intrinsically disordered proteins. *Current Opinion in Structural Biology* **19**, 23-30 (2009).
104. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
105. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **50**, D439-D444 (2022).
106. Wilson, C. J., Choy, W.-Y. & Karttunen, M. AlphaFold2: A Role for Disordered Protein/Region Prediction? *IJMS* **23**, 4591 (2022).
107. Tunyasuvunakool, K. *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590-596 (2021).
108. Kirschner, K. N. *et al.* GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J Comput Chem* **29**, 622-655 (2008).
109. Grant, O. C. *et al.* Gly-Spec: a webtool for predicting glycan specificity by integrating glycan array screening data and 3D structure. *Glycobiology* **26**, 1027-1028 (2016).
110. Grant, O. C. *et al.* Combining 3D structure with glycan array data provides insight into the origin of glycan specificity. *Glycobiology* **26**, 772-783 (2016).

111. Canis, K. *et al.* In-depth comparison of N-glycosylation of human plasma-derived factor VIII and different recombinant products: from structure to clinical implications. *Journal of Thrombosis and Haemostasis* **16**, 1592-1603 (2018).
112. Delignat, S. *et al.* Removal of Mannose-Ending Glycan at Asn2118 Abrogates FVIII Presentation by Human Monocyte-Derived Dendritic Cells. *Front. Immunol.* **11**, 393 (2020).
113. Kannicht, C., Kröning, M., Solecka-Witulska, B., Kohla, G. & Rosenlöcher, J. Comparative N-Glycosylation Analysis of the Fc Portions of a Chimeric Human Coagulation Factor VIII and Immunoglobulin G1. *Bioengineering* **4**, 44 (2017).
114. He, B. *et al.* Stable expression of native Coagulation factor VIII using the 2A self-processing sequence and furin cleavage site. *Thrombosis Research* **128**, e148-e153 (2011).
115. Elkouby, L., Davidson, R. J., Margaritis, P., High, K. A. & Sabatino, D. E. Improved Factor VIII Expression with the Introduction of a PACE-Furin Variant into Codon-Optimized Human Factor VIII. *Blood* **124**, 3498-3498 (2014).
116. Nguyen, G. N. *et al.* Novel factor VIII variants with a modified furin cleavage site improve the efficacy of gene therapy for hemophilia A. *Journal of Thrombosis and Haemostasis* **15**, 110-121 (2017).
117. Nguyen, G. N. *et al.* Human Factor VIII Variants with Diminished PACE-Furin Cleavage Exhibit Increased Biological Activity and Therapeutic Efficacy. *Blood* **124**, 104-104 (2014).
118. Siner, J. I. *et al.* Unexpected Role of PACE/Furin Cleavage Site in FVIII Biology: Implications for Hemophilia a Therapy. *Blood* **124**, 105-105 (2014).
119. Krieger, E., Koraimann, G. & Vriend, G. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins* **47**, 393-402 (2002).
120. Land, H. & Humble, M. S. YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. in *Protein Engineering* (eds. Bornscheuer, U. T. & Höhne, M.) vol. 1685 43-67 (Springer New York, New York, NY, 2018).
121. Ngo, J. C. K., Huang, M., Roth, D. A., Furie, B. C. & Furie, B. Crystal Structure of Human Factor VIII: Implications for the Formation of the Factor IXa-Factor VIIIa Complex. *Structure* **16**, 597-606 (2008).
122. Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Research* **32**, W590-W594 (2004).
123. Chen, S. & Pellequer, J. Identification of Functionally Important Residues in Proteins Using Comparative Models. *CMC* **11**, 595-605 (2004).
124. Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272-272 (1996).
125. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Marius Clore, G. The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance* **160**, 65-73 (2003).
126. Schrödinger, LLC. The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, Version 1.8. (2015).
127. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
128. Ay, C. *et al.* Classification of recombinant factor VIII products and implications for clinical practice: A systematic literature review. *Haemophilia* **30**, 577-588 (2024).
129. Schiavoni, M. *et al.* Status of Recombinant Factor VIII Concentrate Treatment for Hemophilia A in Italy: Characteristics and Clinical Benefits. *Front. Med.* **6**, 261 (2019).
130. Cohen, H. *et al.* Calcium Prevents Enhanced Degradation of Factor VIII in the Condition of Motion. *Biology* **12**, 1388 (2023).
131. Mikaelsson, M., Forsman, N. & Oswaldsson, U. Human factor VIII: a calcium-linked protein complex. *Blood* **62**, 1006-1015 (1983).
132. Switzer, M. E. & McKee, P. A. Some Effects of Calcium on the Activation of Human Factor VIII/Von Willebrand Factor Protein by Thrombin. *J. Clin. Invest.* **60**, 819-828 (1977).
133. Micsonai, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Research* **46**, W315-W322 (2018).
134. Miles, A. J., Ramalli, S. G. & Wallace, B. A. DichroWeb , a website for calculating protein secondary structure from circular dichroism spectroscopic data. *Protein Science* **31**, 37-46 (2022).
135. Whitmore, L. & Wallace, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* **32**, W668-W673 (2004).
136. Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **89**, 392-400 (2008).
137. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. *Open Physics* **10**, (2012).
138. Chen, S. W. & Pellequer, J.-L. DeStripe: frequency-based algorithm for removing stripe noises from AFM images. *BMC Struct Biol* **11**, 7 (2011).

139. C.P.van Zundert, G., M.J.J. Bonvin, A., & Bijvoet Center for Biomolecular Research, Faculty of Science—Chemistry, Utrecht University, Utrecht, the Netherlands. Fast and sensitive rigid-body fitting into cryo-EM density maps with PowerFit. *AIMS Biophysics* **2**, 73-87 (2015).
140. Afonine, P. V. *et al.* New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol* **74**, 814-840 (2018).
141. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis: UCSF ChimeraX Visualization System. *Protein Science* **27**, 14-25 (2018).
142. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *Journal of Structural Biology* **157**, 281-287 (2007).
143. Chaves, R. C. & Pellequer, J.-L. DockAFM: benchmarking protein structures by docking under AFM topographs. *Bioinformatics* **29**, 3230-3231 (2013).
144. Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Science* **27**, 112-128 (2018).
145. Bank, R. E. & Holst, M. A New Paradigm for Parallel Adaptive Meshing Algorithms. *SIAM Rev.* **45**, 291-323 (2003).
146. Sanner, M. F., Olson, A. J. & Spehner, J.-C. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* **38**, 305-320 (1996).
147. Walter, J. D. *et al.* Characterization and Solution Structure of the Factor VIII C2 Domain in a Ternary Complex with Classical and Non-classical Inhibitor Antibodies. *Journal of Biological Chemistry* **288**, 9905-9914 (2013).
148. Wuerth, M. E., Cragerud, R. K. & Clint Spiegel, P. Structure of the Human Factor VIII C2 Domain in Complex with the 3E6 Inhibitory Antibody. *Sci Rep* **5**, 17216 (2015).
149. Spiegel, P. C., Murphy, P. & Stoddard, B. L. Surface-exposed Hemophilic Mutations across the Factor VIII C2 Domain Have Variable Effects on Stability and Binding Activities. *Journal of Biological Chemistry* **279**, 53691-53698 (2004).
150. Coxon, C. H. *et al.* Characterisation and application of recombinant FVIII-neutralising antibodies from haemophilia A inhibitor patients. *Br J Haematol* **193**, 976-987 (2021).
151. Dimitrov, J. D. *et al.* A human FVIII inhibitor modulates FVIII surface electrostatics at a VWF-binding site distant from its epitope. *Journal of Thrombosis and Haemostasis* **8**, 1524-1531 (2010).
152. Pittman, D. D. & Kaufman, R. J. Proteolytic requirements for thrombin activation of anti-hemophilic factor (factor VIII). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2429-2433 (1988).
153. Goodeve, A., Williams, I., Bray, G., Peake, I., & for the Recombinate^{TM} PUP Study Group. Relationship between Factor VIII Mutation Type and Inhibitor Development in a Cohort of Previously Untreated Patients Treated with Recombinant Factor VIII (Recombinate^{TM}). *Thromb Haemost* **83**, 844-848 (2000).
154. Hart, D. P. FVIII Immunogenicity—Bioinformatic Approaches to Evaluate Inhibitor Risk in Non-severe Hemophilia A. *Front. Immunol.* **11**, 1498 (2020).
155. Doshi, B. S., Gangadharan, B., Doering, C. B. & Meeks, S. L. Potentiation of Thrombin Generation in Hemophilia A Plasma by Coagulation Factor VIII and Characterization of Antibody-Specific Inhibition. *PLoS ONE* **7**, e48172 (2012).
156. Witmer, C. & Young, G. Factor VIII inhibitors in hemophilia A: rationale and latest evidence. *Therapeutic Advances in Hematology* **4**, 59-72 (2013).
157. Tabriznia-Tabrizi, S., Gholampour, M. & Mansouritorghabeh, H. A close insight to factor VIII inhibitor in the congenital hemophilia A. *Expert Review of Hematology* **9**, 903-913 (2016).
158. Repessé, Y. *et al.* Factor VIII (FVIII) gene mutations in 120 patients with hemophilia A: detection of 26 novel mutations and correlation with FVIII inhibitor development. *Journal of Thrombosis and Haemostasis* **5**, 1469-1476 (2007).
159. Pahari, S. *et al.* SAAMBE-3D: Predicting Effect of Mutations on Protein-Protein Interactions. *IJMS* **21**, 2563 (2020).
160. Moreira, I. S. *et al.* SpotOn: High Accuracy Identification of Protein-Protein Interface Hot-Spots. *Sci Rep* **7**, 8007 (2017).
161. Bonazza, K. *et al.* Ca2+ concentration-dependent conformational change of FVIII B-domain observed by atomic force microscopy. *Anal Bioanal Chem* **407**, 6051-6056 (2015).
162. Parmenter, C. D. J., Cane, M. C., Zhang, R. & Stoilova-McPhie, S. Cryo-electron microscopy of coagulation Factor VIII bound to lipid nanotubes. *Biochemical and Biophysical Research Communications* **366**, 288-293 (2008).
163. Ruben, E. A., Rau, M. J., Fitzpatrick, J. A. J. & Di Cera, E. Cryo-EM structures of human coagulation factors V and Va. *Blood* **137**, 3137-3144 (2021).

164. Zhang, Y., Srivastava, V. & Zhang, B. Mammalian cargo receptors for endoplasmic reticulum-to-Golgi transport: mechanisms and interactions. *Biochemical Society Transactions* **51**, 971-981 (2023).
165. Yakovleva, E. & Zhang, B. Clinical, Laboratory, Molecular, and Reproductive Aspects of Combined Deficiency of Factors V and VIII. *Semin Thromb Hemost* s-0044-1789019 (2024) doi:10.1055/s-0044-1789019.
166. Vander Kooi, A. *et al.* Influence of N-glycosylation in the A and C domains on the immunogenicity of factor VIII. *Blood Advances* **6**, 4271-4282 (2022).
167. Lai, J. D. & Lillicrap, D. Factor VIII inhibitors: Advances in basic and translational science. *Int J Lab Hem* **39**, 6-13 (2017).
168. Qu, J. *et al.* Comparative glycosylation mapping of plasma-derived and recombinant human factor VIII. *PLoS ONE* **15**, e0233576 (2020).
169. Turecek, P. L. *et al.* BAX 855, a PEGylated rFVIII product with prolonged half-life. Development, functional and structural characterisation. *Hamostaseologie* **32 Suppl 1**, S29-38 (2012).