Received: 7 March 2024 | Revised: 10 June 2024 | Accepted: 5 July 2024

https://doi.org/10.1016/j.jtha.2024.07.008

ORIGINAL ARTICLE

An integrated multitool analysis contributes elements to interpreting unclassified factor IX missense variants associated with hemophilia B

Monica Sacco[1] | Maria Francesca Testa[2] | Antonietta Ferretti[3] | Maria Basso[3] | Stefano Lancellotti[3] | Maira Tardugno[1] | Leonardo Di Gennaro[3] | Paola Concolino[4] | Angelo Minucci[4] | Claudia Spoliti[1] | Alessio Branchini[2] | Raimondo De Cristofaro[1,3]

[1]Department of Translational Medicine and Surgery, Catholic University of the Sacred Heart, Rome, Italy

[2]Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy

[3]Center for Hemorrhagic and Thrombotic Diseases, Foundation University Hospital "A. Gemelli" IRCCS, Rome, Italy

[4]Molecular and Genomic Diagnostics Unit, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

Correspondence
Raimondo De Cristofaro, Department of Translational Medicine and Surgery, Catholic University of the Sacred Heart; Largo Agostino Gemelli 8, 0168, Rome, Italy.
Email: raimondo.decristofaro@unicatt.it

Alessio Branchini, Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, via Fossato di Mortara 74, 44121, Ferrara, Italy.
Email: brnlss@unife.it

Abstract

Background: Dissection of genotype–phenotype relationships in hemophilia B (HB) is particularly relevant for challenging (mild HB) or for HB-associated but unclassified factor (F)IX missense variants.

Objective: To contribute elements to interpret unclassified HB-associated FIX missense variants by a multiple-level approach upon identification of a reported, but uncharacterized, FIX missense variant associated with mild HB.

Methods: Molecular modeling of wild-type and V92A FIX variants, expression studies in HEK293 cells with evaluation of protein (ELISA, western blotting) and activity (activated partial thromboplastin time-based/chromogenic assays) levels after recombinant expression, and multiple prediction tools.

Results: The *F9*(NM_000133.4):c.275T>C (p.V92A) variant was found in a mild HB patient (antigen, 45.4 U/dL; coagulant activity, 23.6 IU/dL; specific activity, 0.52). Newly generated molecular models showed alterations in Gla/EGF1–EGF2 domain conformation impacting Ca^{++} affinity and protein–protein interactions with activated factor XI (FXIa). Multitool analysis indicated a moderate impact on protein structure/function of the valine-to-alanine substitution, in accordance with patient and modeling data. Expression studies on the V92A variant showed a specific activity (0.49 ± 0.07; wild-type, 1.0 ± 0.1) recapitulating that of the natural variant, and pointed toward a moderate activation impairment as the main determinant underlying the p.V92A defect. The validated multitool approach, integrated with evidence-based data, was challenged on a panel (n = 9) of unclassified FIX missense variants, which resulted in inferred protein (secretion/function) outputs and HB severity.

Manuscript handled by: Karl C. Desch

Final decision: Karl C. Desch, 5 July 2024

Monica Sacco, Maria Francesca Testa, and Antonietta Ferretti contributed equally to this study.

© 2024 The Author(s). Published by Elsevier Inc. on behalf of International Society on Thrombosis and Haemostasis. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Conclusion: The rational integration of multitool and multiparameter analyses contributed elements to interpret genotype/phenotype relationships of unclassified FIX missense variants, with implications for diagnosis, management, and treatment of HB patients, and potentially translatable into other human disorders.

KEYWORDS
genotype–phenotype association, hemophilia B, missense mutation, molecular models, prediction tools

1 | INTRODUCTION

Hemophilia B (HB, OMIM #306900), a recessive X-linked bleeding disorder characterized by partial deficiency or absence of coagulation factor (F)IX [1], is caused by *F9* gene (OMIM #300746) alterations. Based on the different HB severity degrees, namely severe (<1%), moderate (1%-5%), and mild (5%-40%) [2], missense mutations, the most frequent point mutation type [3], are associated with the highest phenotype heterogeneity. Although severe HB, associated with spontaneous bleeding, and moderate HB, with bleeding primarily occurring after injury or minor surgery, phenotypes are well-defined, mild HB may be the most challenging from a clinical standpoint, as subjects may unexpectedly undergo severe hemorrhages upon surgical interventions or major trauma [2–5].

These elements may result in a lack of classification for a proportion of HB-associated FIX variants, whose characterization/dissection would help deepen the knowledge on the genotype–phenotype relationships in HB for challenging (ie, mild HB) or unclassified FIX variants [6], as well as on FIX structural/functional relationships in terms of interactions with other effectors within the coagulation cascade. In this view, the availability of models of whole FIX with included relevant domains (ie, Gla and EGF1-2) would provide a tool for structural inspections.

In this study, we took advantage of the structural and functional characterization of the unclassified *F9* p.V92A missense variant, showing a decreased activity/antigen ratio mainly due to an activation defect, as paradigmatic reference to provide FIX structural insights through a new molecular model of zymogen FIX, currently not available, and propose a classification of uncategorized FIX missense variants by an extensive analysis based on multiparameter prediction. Our multiple-level approach contributes elements to interpret reported, but unclassified, HB-associated FIX missense variants, supporting computational approaches as tools to integrate clinical, genetic, and biochemical analyses, with implications for diagnosis, management, and treatment of unclassified HB patients.

2 | MATERIALS AND METHODS

2.1 | Patient

A 45-year-old male patient was referred from the emergency room for ileo–psoas hematoma under warfarin oral anticoagulant therapy (OAT) due to mechanical aortic valve replacement in 2003. The patient had a history of Takayasu disease, hypertension, and previous psoas hematomas since OAT started. He reported, without grounds, the onset of pain in the left groin, radiating into the ipsilateral iliac fossa and along the thigh, progressively worsening. For the painful symptoms, he underwent ultrasound examination, which found a vast hematoma in the abdominal wall. In March 2023, the patient underwent aortic valve replacement, associated with the replacement of the aortic bulb and ascending aorta according to the Bentall method. The surgical intervention was carried out while the patient was under prophylaxis with albutrepenonacog alfa (Idelvion, 2000 U) infusions, with FIX coagulant levels maintained at >70%. The main laboratory parameters of the patient after cessation of warfarin therapy are listed in Supplementary Table S1.

The patient showed constant alteration of the activated partial thromboplastin time (APTT) ratio (1.51-1.55), suggesting a defect in the intrinsic coagulation pathway. Analysis of patient plasma by polyclonal ELISA and APTT assays showed FIX:Ag (45.4%) and FIX coagulant (23.6%), with a specific activity (0.52; normal value, 1.0), compatible with mild HB and a secretion and/or functional defect. The thrombin generation assay, performed by using the BleedScreen reagent (Stago) in a ST Genesia platform (Stago) showed reduced parameters compatible with impaired thrombin generation capacity compared with control plasma (Supplementary Figure S1), further supporting mild HB. Of note, the patient was unaware of his mild HB condition. The study was carried out in accordance with the Declaration of Helsinki, and written informed consent was obtained.

2.2 | Nomenclature

Amino acid positions are numbered according to the Human Genome Variation Society nomenclature [7,8], with numbering starting at the AUG (codon 1) translation initiation codon.

2.3 | Genetic analysis

Massive parallel sequencing was carried out using the capture-based Next Generation Sequencing Clinical Exome Solution v3 kit (SOPHIA GENETICS) on an Illumina MiSeq instrument (Illumina). We

designed a virtual panel of 87 genes associated with bleeding thrombotic and platelet disorder according to International Society on Thrombosis and Haemostasis (ISTH) 2023 guidelines. Sequencing data were analyzed by Sophia DDM software v.4.2 (SOPHIA GENETICS).

2.4 | Modeling of the p.V92A missense variant and prediction of residues coordinating Ca^{++}

Models of wild-type (WT-) and p.V92A (V92A-) FIX variants (sequence 47-461) were obtained by using the I-TASSER (Iterative ThreadingASSEmbly Refinement) threading modeling server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/), as previously detailed [9,10]. After having obtained the best models, each model was analyzed with the FG-MD program [11]. This software first identifies analogous fragments from the protein data bank (PDB) by the structural alignment program TM-align [12]. Spatial restraints extracted from the fragments are then used to reshape the funnel of the MD energy landscape and guide the molecular dynamics conformational sampling. FG-MD aims at refining the initial models closer to the native structure. It can also improve the local geometry of the structures by removing the steric clashes and improving the torsion angle and the hydrogen-binding networks.

Prediction of calcium ion interaction site was attained through the MIB2 program, which provides the energetics of coordinating side chains of a protein and scores their potential [13].

2.5 | Modeling of FIX-FXIa interactions and relative energetics

The X-ray solved crystal structure of FXIa in complex with the inhibitor methyl [(5E,8S)-8-[(6R)-6-(3-chlorophenyl)-2-oxo-1,3-oxazinan-3-yl]-2-oxo-1,3,4,7,8,10-hexahydro-2H-12,9-(azeno)-1,10-benzodiazacyclotetradecin-15-yl]carbamate (PDB file 5QQO) was used to investigate the interaction with both WT-FIX and V92A-FIX. In the molecular modeling investigation, the inhibitor was removed by using the PyMOL program. The docking complexes of WT-FIX and V92A-FIX refined models with FXIa (PDB file 5QQO) were investigated through the ClusPro program [14], which employs Fast Fourier Transform (FFT) for energy evaluation to investigate protein–protein docking *in silico* [15]. The FFT-based algorithm enables global docking without any *a priori* information on the structure of the adduct. The domains involved in the docking process and the relative contribution of hydrogen and hydrophobic energies were also analyzed for each complex by using the PDBsum platform (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/). To validate and corroborate the results obtained with the ClusPro program in modeling the FXIa-FIX complex, the 2 proteins were used alternatively as ligand and receptor, obtaining in all cases the same result. The graphical rendering of the various models was performed using the PyMOL program.

Prediction of binding affinity either of WT-FIX or V92A-FIX with FXIa models was calculated through the PRODIGY program, which calculates the interaction energy of protein–protein complexes from their 3D structure, taking into consideration hydrogen and hydrophobic contributions to the binding energy [16].

2.6 | Expression studies on the p.V92A variant

The p.V92A variant was mimicked by site-directed mutagenesis (Agilent Technologies) of *F9* cDNA (reference sequences, GenBank: NG_007994.1, NM_000133.4, and NP_000124.1), cloned in the pCDNA3 vector, through the 5′-GAAGCAGTATGCTGATGGAGATC-3′ oligonucleotide (modified nucleotide underlined; reverse primer complementary). The resulting plasmid was validated by sequencing. Mutagenesis and expression studies, carried out in HEK293 cells, were as described [17].

2.7 | Evaluation of FIX protein and activity levels

FIX protein and activity levels on plasma samples or cell medium, as well as western blotting analysis, were evaluated essentially as described [17]. Briefly, FIX protein levels were evaluated by polyclonal anti-human FIX ELISA (FIX-EIA, Affinity Biologicals), and protein forms in media and cell lysates by western blotting analysis with polyclonal goat anti-human FIX (#GAFIX-AP; Affinity Biologicals) and polyclonal anti-goat horseradish peroxidase(HRP)-conjugated (#A50-101P; Bethyl Laboratories) antibodies, as well as by polyclonal rabbit anti-BiP (#3183; Cell Signaling Technology) and monoclonal mouse anti-GAPDH (#MA5-15738; Invitrogen) primary antibodies followed by detection with HRP-conjugated polyclonal goat anti-rabbit (#P0448; DAKO, Agilent) and polyclonal goat anti-mouse (#P0447; DAKO, Agilent) antibodies. Time-course experiments were performed by incubating FIX variants in medium with plasma-derived FXIa (Haematologic Technologies Inc) diluted in Reaction buffer (Hepes, 20 mM; NaCl, 150 mM; PEG-8000, 0.1%; $CaCl_2$, 5 mM; pH, 7.4), or by adapting a commercially available chromogenic assay (HYPHEN BioMed; Aniara Diagnostica). Protein forms were evaluated by anti-FIX antibodies as described above, or by polyclonal sheep anti-human FX (#CL20047AP, Cedarlane) and HRP-conjugated polyclonal rabbit anti-sheep (#P0163, DAKO) antibodies [18]. Blotting images were acquired on the ChemiDoc instrument and analyzed by the Image Laboratory Software version 4.0 (Bio-Rad).

2.8 | Multitool analysis of unclassified FIX missense variants

Genetic and epidemiologic information on FIX missense variants were retrieved by inspection of the European Association of Hemophilia and Allied Disorders (EAHAD, https://f9-db.eahad.org/index.php) [19,20], CDC Hemophilia Project (CHBMP, https://www.cdc.gov/ncbddd/hemophilia/champs.html) [21] and the Human Genome Mutation (HMGD, http://www.hgmd.cf.ac.uk/ac/all.php) [22] databases.

The impact of amino acid substitutions at the selected FIX positions was predicted by multiple tools (n = 15; Variant Effector Predictor, https://www.ensembl.org/Homo_sapiens/Tools/VEP) to counterbalance the predicted output stemming from each tool, and thus improve reliability of final prediction and of the resulting protein output. Scores and cutoff values were retrieved by inspection of reference values provided for each tool, whose brief description is provided as Supplementary Method. The prediction output was indicated by 0 (tolerated), 0.5 (intermediate) and 1 (damaging), which resulted in a calculated tolerance score (ranging from 0 to 1) obtained by the ratio between the number of 0s, recurring after prediction with the 15 tools, on the total of prediction outputs. Prediction of protein stability was evaluated by the deep learning-based prediction tool DDMut [23], which provides ΔΔG values (kcal/mol; hereafter referred as $\Delta\Delta G_{stab}$), reflecting a stabilizing ($\Delta\Delta G \geq 0$ kcal/mol) or destabilizing ($\Delta\Delta G < 0$ kcal/mol) effects of the amino acid substitution.

2.9 | Statistical analysis

Data were analyzed by unpaired t-tests, with the value of $P < .05$ considered as statistically significant.

3 | RESULTS

3.1 | Genetic analysis

Next generation sequencing analysis revealed the *F9*(NM_000133.4):c.275T>C variant, causing the p.V92A substitution at the interface between the Gla and EGF1 domains of the FIX molecule. This variant was previously reported in the EAHAD, CHBMP, and HMGD databases, but, to date, no phenotypic information is available in terms of FIX levels or HB severity. In addition, no data regarding the frequency in general population were obtained from gnomAD v4.0.0 database (https://gnomad.broadinstitute.org/). Analysis with VarSome software (http://varsome.com) classified the p.V92A as "Variant of Uncertain Significance".

3.2 | Generation and validation of the V92A-FIX molecular model

Prompted by the availability of coevolution-based contact prediction [24–26] and deep neural-network learning [27] techniques, we performed molecular modeling on the V92A-FIX variant to provide insights into the impact of this amino acid substitution on FIX structural and functional properties.

Valine 92 of human FIX is localized at the interface between Gla and EGF1 domains of the protein. Unfortunately, no X-ray or NMR data concerning zymogen FIX are available. Thus, we took advantage of the X-ray-solved structure of FIXa (PDB file 6MV4) to perform molecular modeling of the remaining FIX domains (Gla domain, EGF1, EGF2, and activation peptide sequence). The 2D NMR-solved structure of EGF1 of human FIX (PDB file 1IXA) was used as overall molecular template of the EGF domains [28].

I-TASSER generated acceptable structure predictions for both WT-FIX and V92A-FIX variants, as indicated by the superimposed structures (Figure 1A). The best model was selected based on the convergence of the structures with the lowest intermolecular energy. The TM-score of the predicted model was found equal to 0.42 ± 0.14 for both the WT-FIX and V92A-FIX models. An element that confirmed the acceptable goodness of the molecular models of both zymogen WT-FIX and V92A-FIX was represented by the superimposable catalytic domain with the X-ray crystal structure of FIXa (PDB file 6MV4) (Figure 1B, C) [29]. A closer inspection of the models revealed subtle but significant topological differences for the catalytic triad (His267 [c57], Asp315 [c102], Ser411 [c195]), whereby in the mutant some atoms (His57 Nδ1/Asp102 Oδ1) engaged in the proton transfer of the catalytic cycle are up to 0.7 Å more distant than in the WT protein (Figure 1D). This effect may substantially slow down the nucleophilic attack by Ser 411 of its substrate [30].

The newly generated and validated V92A-FIX model, resolving FIX regions absent in other structures currently available, provided first insights into the effects of the valine-to-alanine substitution on the properties of the catalytic domain in the p.V92A missense variant.

3.3 | Comparative analysis of WT-FIX and V92A-FIX structures

To dissect the local impact of alanine substituting the original valine residue, we performed comparative structural analysis of the best models of WT-FIX and V92A-FIX, which showed an RMSD value (13.3 ± 4.1) reflecting a significant difference in the 3D configuration of the sequences due to the presence (WT-FIX; Figure 2A) or absence (V92A-FIX; Figure 2B) of a methyl group. The main differences concern the 54-65 region (Gla domain) and mostly the EGF1-2 domains, which, compared with the WT form, are particularly disordered. Notably, previous studies showed that the substitution of D93 with a glycine residue was also identified as responsible for a moderate form of HB (FIX Alabama), as it causes a conformational change in the EGF1-2 domains that propagates to the serine protease domain [31]. Thus, the absence of a methyl group at position 92 (alanine vs valine) causes a perturbation of hydrogen bonds occurring between V92 and W88 (between NH of valine and CO of W88), as shown in Figure 2A, B. Hence, the lack of methyl group in A92 causes a rearrangement of hydrogen bonds in the W88-K89 region. The NH group of A92 interacts with high energy (distance <3Å) with the oxygen atom of the CO groups of W88 and K89. The enhanced interaction between A92 and W88/K89 residues causes a sort of "retraction" of the C-terminal portion of the Gla domain, which alters the tertiary conformation of EGF1-2 domains (Figure 1A and Figure 2A, B). EGF1-2 domains in FIX act as spacers that position the active site of FIXa at optimal distance from the membrane to allow interaction with its cofactor (FVIIIa) and substrate (FX) [32–34]. Moreover, D110 in the N-terminal EGF1

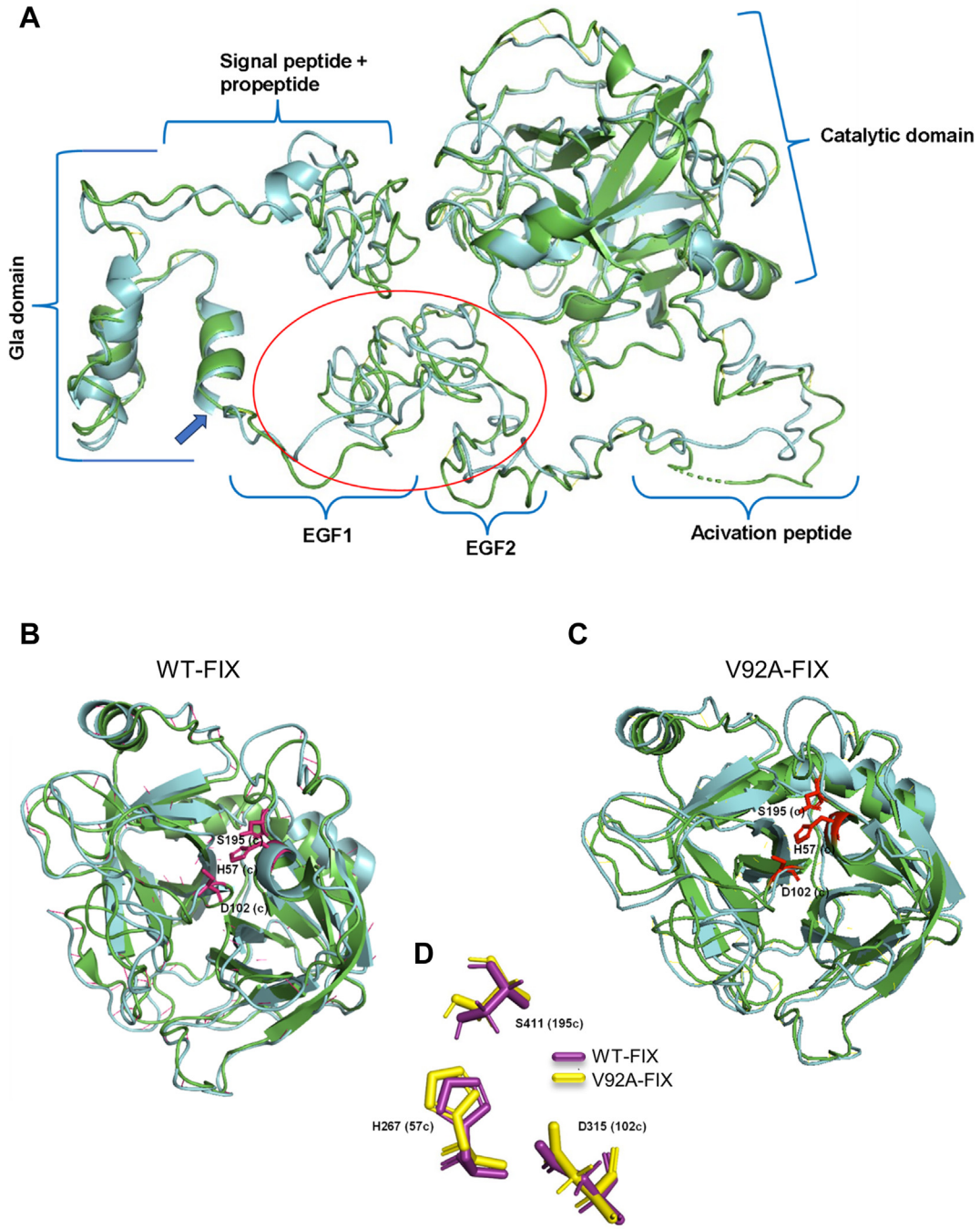

FIGURE 1 Modeling of wild-type (WT)-factor(F)IX and V92A-FIX. (A) Superposition of the molecular models of zymogen WT-FIX (green cartoon) and V92A-FIX (cyan cartoon) variants. The different domains of the FIX protein are indicated in the figure, whereas the red circle highlights the region mostly perturbed by the p.V92A missense mutation involving FIX EGF1-2 domains. The blue arrow indicates the localization of V92. (B) Superposition of the catalytic domain of WT-FIX (PDB file 6MV4), shown as a green cartoon, and the same domain modeled by I-TASSER (cyan). (C) Superimposition of V92A-FIX (cyan cartoon) as in (B). The catalytic domain, shown in purple and red sticks for WT and V92A, respectively, is also shown and numbered according to the chymotrypsin(ogen) numbering system. (D) Topological differences in side chains of the catalytic triad residues of WT-FIX (purple) and V92A-FIX (yellow).

domain, together with D111 and N113, is engaged in coordinating 1 Ca^{++}, which allosterically promotes FIX coagulant activity [35]. The 3 residues D110, D111, and N113 were found to be characterized by the highest score value (3.49) for Ca^{++} binding, according to the result obtained with the MIB2 program in being responsible for the coordination of a calcium ion (Figure 3A). This prediction is in accordance with previous X-ray structure of FIXa [36]. At variance, the coordinating binding site constituted by D110 and D111 only is substantially

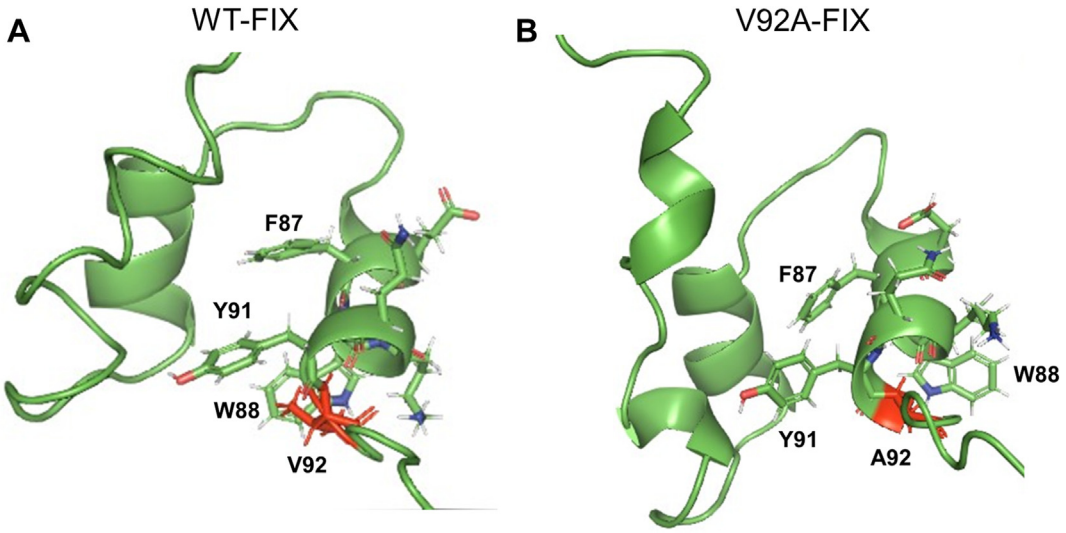

FIGURE 2 Model of local impact of the V92A amino acid substitution. (A) Hydrogen bonds between A92 and oxygen atoms of W88 and K89, whose distance is typical of high energy interactions (distance: 2.49 and 2.33 Å, respectively) in wild-type (WT)-factor (F)IX. (B) Hydrogen bond between N atom of V92 with oxygen of W88 (yellow dashed line) in V92-FIX. The distance between the hydrogen of V92 and acceptor of Y88 is 2.07 Å.

lost in the V92A-FIX variant, being characterized by a significantly lower score value (1.19) predicted by the MIB2 program.

These findings contributed to dissecting the local perturbations associated with the V92A amino acid substitution.

These data provided further insights into the altered interaction with a key (feedback) FIX activator (FXIa), with impact on structural and functional properties of the p.V92A FIX missense variant and in line with its reduced coagulant capacity.

3.4 | Interaction of WT-FIX and V92A-FIX with FXIa

The refinement of WT-FIX and the V92A-FIX models allowed us to evaluate their interaction with FXIa, with docking complexes investigated through the ClusPro program, which employs an FFT-based computational strategy for energy evaluation in protein–protein docking [15]. In FFT-based methods, one of the proteins is placed at the origin of the coordinate system on a fixed grid, the interacting protein is instead placed on a movable grid, and the interaction energy is computed as a sum of some correlation functions. With this computational strategy, these energy functions can be simultaneously evaluated for all transitions using FFT, and only rotations need to be explicitly considered. This results in the ability to exhaustively sample billions of conformations of the 2 interacting proteins, obtaining energy values at each grid point.

The models of FXIa monomer interaction with WT-FIX and V92A-FIX are shown in Figure 3B and Figure 3C, respectively. The main interaction parameters obtained with the program PDBsum are listed in Table 1. Notably, the presence of A92 in the FIX mutant induces a conformational change in the 88-93 segment, so that the hydrogen bonds with R170, R172, and Q221 residues in FXIa are lost (Figure 3B). These changes are reflected by a loss of energy, as revealed by the application of the Prodigy program, which provided, *in silico* (T = 37 °C), K_d values corresponding to 4.11×10^{-11} M and 2×10^{-8} M for WT-FIX and V92A-FIX, respectively.

3.5 | Multitool analysis and expression studies on the V92A missense variant

Molecular modeling provided key information on structural properties of WT-FIX as well as of the p.V92A variant, including the differential engagement of interactions with a crucial coagulation factor such as FXIa.

To provide elements to further dissect the functional impact of the valine-to-alanine substitution, we predicted the effects of each possible amino acid substitution stemming from the original GTT triplet by a panel of multiple prediction tools, as well as by assessing octanol-water partition coefficient ($\log K_{ow}$) and protein stability ($\Delta\Delta G_{stab}$) by DDmut (Figure 4A). The valine-to-alanine substitution was predicted as tolerated by 6/15 tools, among those substitutions mostly predicted as damaging (4/15 tools, Val>Asp/Gly) or tolerated (10/15, Val>Ile/Leu), corresponding to intermediate $\log K_{ow}$ and $\Delta\Delta G_{stab}$ values (Figure 4A). Interestingly, comparison of the hydrophobic properties of amino acids ($r = 0.96$; $P = .01$; Figure 4B, left panel) or the calculated tolerance score ($r = 0.99$; $P = .0001$; Figure 4B, right panel) with protein stability showed significant correlations. Overall, this analysis predicted the valine-to-alanine substitution as exerting a moderate impact on protein properties.

We recombinantly expressed the FIX-V92A variant in HEK293 cells to validate the predicted impact and properties associated with the V92A substitution, as well as to provide experimental evidence

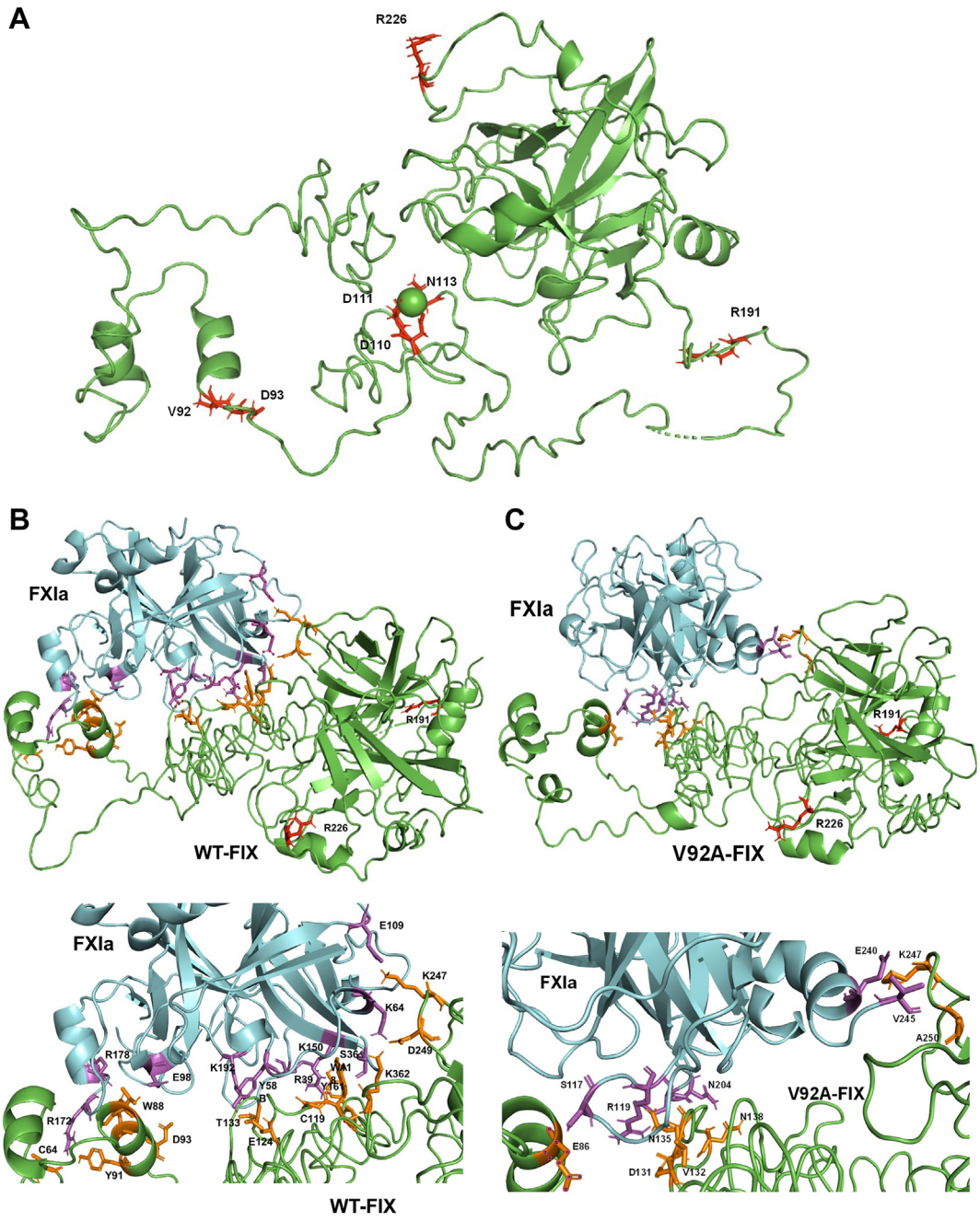

FIGURE 3 Analysis of interaction between wild-type (WT)-factor(F)IX or V92A-FIX with FXIa. (A) Model of WT-FIX. The calcium ion (green ball) is coordinated by D110, D111, and N113 side chain residues. R191 and R216 are the 2 arginine residues in the FIX activation peptide, whose cleavage by FXIa is responsible for the conversion of FIX to FIXa. D93 is the aspartate mutated to glycine in the mild form of HB referred to as $FIX_{Alabamam}$ [31]. (B) Model of the interaction between FXIa and WT-FIX (B) or V92A-FIX (C) presented as global view (upper panel) showing the 2 arginine residues cleaved by FXIa (R191 and R226), or magnification of the interface (lower panel), where the engaged residues of FIX (orange sticks) and FXIa (magenta sticks) are shown.

on its characterization. The FIX-V92A variant showed lower secreted (110.4 ± 7.3 ng/mL) and intracellular (46.3 ± 5.8 ng/mL) protein levels compared with WT-FIX (secreted, 136.8 ± 14.7 ng/mL; intracellular, 72.8 ± 8.5 ng/mL) (Figure 4C, left panel), thus impacting on FIX protein levels (Figure 4C, right panel). Western blotting analysis did not show relevant differences in both secreted and intracellular protein forms (Figure 4D, upper panels). To shed more light on the processing of the FIX-V92A variant, we evaluated the relative levels of BiP, a chaperone associated with protein folding/misfolding [37]. Analysis of relative intensity between BiP and GAPDH bands revealed a 3-fold lower BiP/GAPDH ratio for the FIX-V92A variant (Figure 4D, lower panels). Evaluation of coagulation times in

TABLE 1 Hydrogen bonds involved in protein–protein interactions between wild-type (WT)- or V92A-factor (F)IX and FXIa.

WT-FIX ↔ FXIa					
Atom number	**Atom name**	**Residue name**	**Residue number**	**Chain**	**Distance (Å)**
623 ↔1600	O ↔ NH2	Cys ↔ Arg	64 ↔ 172	A ↔ B	2.68
836 ↔848	**OG1 ↔ OE1**	**Thr ↔ Glu**	**85 ↔ 98**	**A ↔ B**	**2.92**
872 ↔ 1573	**NE1 ↔ O**	**Trp ↔ Arg**	**88 ↔ 170**	**A ↔ B**	**3.02**
914 ↔ 1597	**OH ↔ NH1**	**Tyr ↔ Arg**	**91 ↔ 172**	**A ↔ B**	**2.71**
914↔ 1600	**OH ↔ NH2**	**Tyr ↔ Arg**	**91 ↔ 172**	**A ↔ B**	**2.81**
931↔ 2062	**OD1↔ NE2**	**Asp ↔ Gln**	**93 ↔ 221**	**A ↔ B**	**2.84**
1145 ↔ 246	O ↔ NH2	Cys ↔ Arg	117 ↔ 39	A ↔ B	2.73
1161 ↔ 240	O ↔ NE	Trp ↔ Arg	118 ↔ 39	A ↔ B	2.90
1161 ↔ 246	O ↔ NH2	Trp ↔ Arg	118 ↔ 39	A ↔ B	2.91
1155 ↔ 205	NE1 ↔ O	Trp ↔ Ser	118 ↔ 36A	A ↔ B	2.78
1166 ↔ 246	SG ↔ NH2	Cys ↔ Arg	119 ↔ 39	A ↔ B	2.75
1212 ↔ 439	OE2 ↔ OH	Glu ↔ Tyr	124 ↔ 58B	A ↔ B	2.87
1295 ↔ 1807	O ↔ NZ	Thr ↔ Lys	133 ↔ 192	A ↔ B	2.76
1560 ↔ 1352	OH ↔ NZ	Tyr ↔ Lys	161 ↔ 150	A ↔ B	2.55
2373 ↔ 493	NZ ↔ O	Lys ↔ Lys	247 ↔ 64	A ↔ B	2.76
2373 ↔ 949	NZ ↔ OE1	Lys ↔ Glu	247 ↔ 109	A ↔ B	2.69
2373 ↔ 950	NZ ↔ OE2	Lys ↔ Glu	247 ↔ 109	A ↔ B	2.56
2393 ↔ 488	OD2 ↔ NZ	Asp ↔ Lys	249 ↔ 64	A ↔ B	2.62
3479 ↔ 202	NZ ↔ OG	Lys ↔ Ser	362 ↔ 36A	A ↔ B	2.62
V92A-FIX ↔ FXIa					
Atom number	**Atom name**	**Residue name**	**Residue number**	**Chain**	**Distance (Å)**
841 ↔ 1016	N ↔ OG	Glu ↔ Ser	86 ↔ 117	A ↔ B	3.13
1276 ↔ 1041	O ↔ NH1	Asp ↔ Arg	131 ↔ 119	A ↔ B	2.63
1284 ↔ 1041	O ↔ NH1	Val ↔ Arg	132 ↔ 119	A ↔ B	2.64
1307 ↔ 1918	ND2 ↔ OE1	Asn ↔ Glu	135 ↔ 205	A ↔ B	2.84
1339 ↔ 1907	OD1 ↔ ND2	Asn ↔ Asn	138 ↔ 204	A ↔ B	2.95
2371 ↔ 2262	NZ ↔ OE1	Lys ↔ Glu	247 ↔ 240	A ↔ B	2.80
2394 ↔ 2313	N ↔ O	Ala ↔ Val	250 ↔ 245	A ↔ B	2.81

F, factor; FXIa, activated factor XI; WT, wild type.

Residues indicated in bold mostly contribute to the interaction of WT-FIX with FXIa.

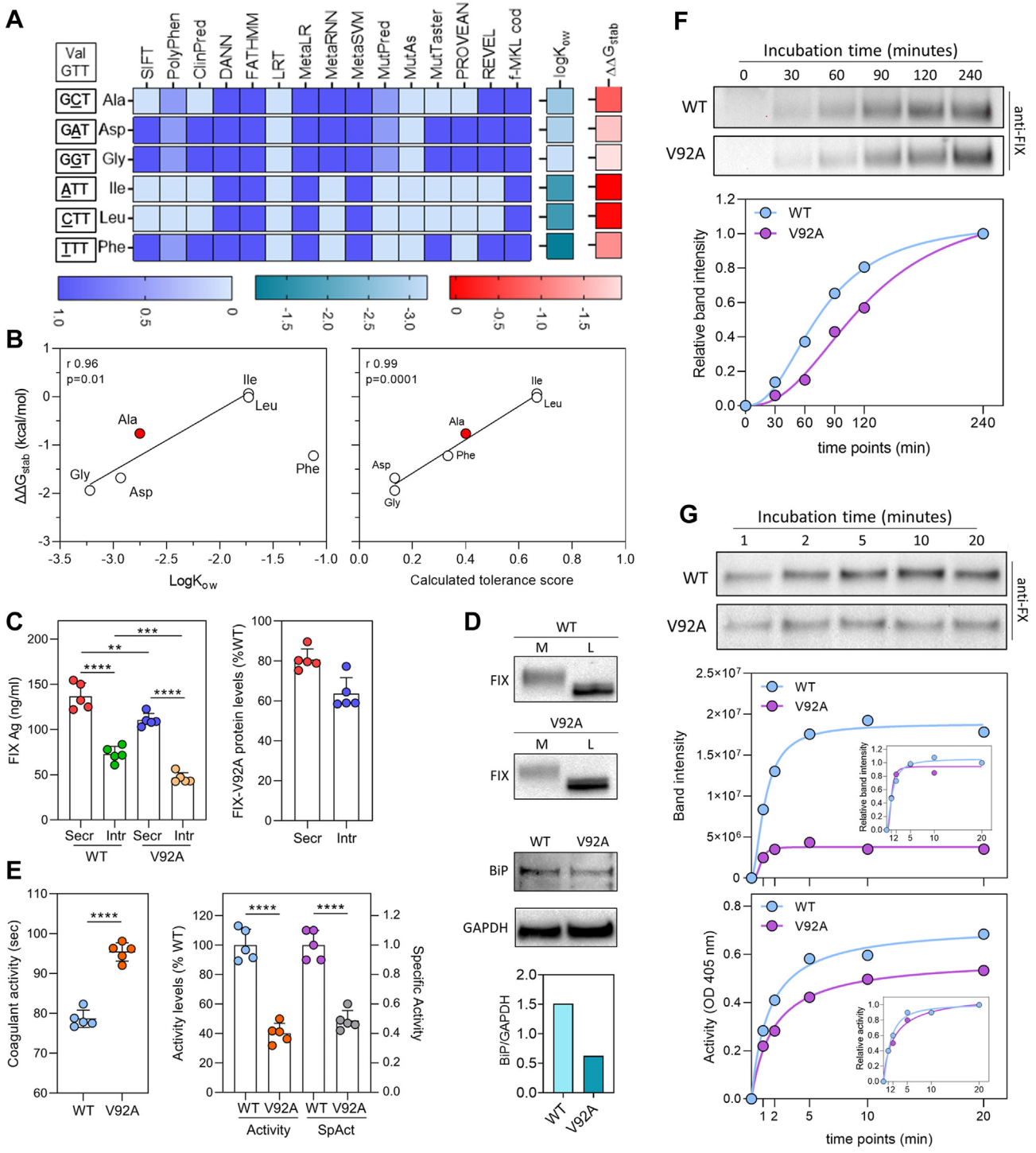

FIGURE 4 Multitool analysis and expression studies on the factor (F)IX-92A variant. (A) Heatmap representing the impact of amino acid substitutions stemming from each possible single-nucleotide change of the original valine-coding GTT codon at position 92. Impact was predicted by multitool analysis (blue shading; bottom legend: 0, tolerated; 0.5, intermediate; 1.0, damaging), octanol-water partition coefficient ($logk_{ow}$, green shading; bottom legend: <−2.0, more hydrophilic; >−2.0, more hydrophobic) and as protein stability by DDmut ($\Delta\Delta G_{stab}$, kcal/mol; red shading; bottom legend: $\Delta\Delta G \geq 0$ kcal/mol, stabilizing effects; $\Delta\Delta G < 0$ kcal/mol, destabilizing effects). Scores and cutoff values were retrieved by inspection of reference values provided for each tool. Codons and the corresponding amino acid, with the original GTT/valine indicated on top, are shown on the left, whereas the exploited tools or parameters are indicated on top of the heatmap. (B) Correlation between the hydrophobic properties of amino acids ($logK_{ow}$; left panel) or the calculated tolerance score (ranging from 0 to 1; right panel) with protein stability ($\Delta\Delta G_{stab}$). The calculated tolerance score derived from the ratio between the number of 0s (indicating a tolerated substitution as described in A) on the total of prediction outputs by the 15 tools. (C) Quantification (polyclonal ELISA) of secreted (Secr; cell medium) and intracellular (Intr; cell lysates) FIX (FIX Ag) of FIX-WT and FIX-V92A variants transiently expressed in HEK293 cells (left panel), and

FIX-depleted plasma (FIX-V92A, 95.4 ± 2.3 seconds; WT-FIX, 78.6 ± 2.2 seconds) (Figure 4E, left panel) indicated a significant impact on coagulant activity (40.2 ± 6.8% of WT-FIX), corresponding to a reduced specific coagulant activity (FIX-V92A, 0.49 ± 0.07; WT-FIX, 1.0 ± 0.1) (Figure 4E, right panel).

We performed time-course experiments in 2 different settings. Western blotting analysis upon incubation of recombinant variants with plasma-derived FXIa (Figure 4F, upper panel) showed that the rate of FXIa-mediated activation was higher for WT-FIX than the FIX-V92A variant, as indicated by analysis of band intensity as a function of the last time point (Figure 4F, lower panel). In the slightly modified chromogenic assay, the FIX-V92A variant showed a lower degree of FIX-dependent FXa activation both in terms of protein forms (band intensity; Figure 4G, upper and middle panels) and chromogenic activity (OD values; Figure 4G, lower panel). Interestingly, the analysis of data as a ratio on the last time point showed a similar relative FXa generation pattern between WT-FIX and FIX-V92A variants over time, as shown, respectively, in insets of middle and lower panels of Figure 4G.

Overall, prediction of the functional impact of the V92A amino acid substitution on FIX structure/function parallels data and information derived from modeling studies and is validated by expression studies, recapitulating patient phenotype, pointing toward an altered FXIa-mediated activation as the main determinant underlying the molecular defect of the p.V92A variant.

3.6 | Multitool analysis and proposed classification for unclassified FIX missense variants

Prompted by the validated data on the p.V92A variant, we extended our analysis to those FIX missense variants being reported as unique (n patients = 1). Inspection of the EAHAD, CHBMP and HGMD databases revealed around 300 unique FIX missense variants, widely distributed in FIX domains, and associated with graded HB severity (Figure 5A). For 30 of them, including the p.V92A, a classification in terms of reported FIX levels or severity is lacking. However, for 20 out of 30 a classification of severity degree can be inferred by overlapping ($n \geq 2$) missense variants with reported severity (Supplementary Table S2). The remaining 9 variants with single (p.P177L, p.Y305N, p.G349E) or unreported (p.N36T, p.Q57E, p.A250V, p.N283D, p.D322Y, p.Y391C) overlapping amino acid substitutions were challenged with our integrated prediction approach.

The multitool analysis predicted a differential impact of each combination of amino acid substitutions stemming from single-nucleotide changes (Supplementary Figure S2 and Supplementary Figure S3, upper heatmap panels), resulting in a wide distribution of calculated tolerance scores, with those of natural variants (red circles) ranging from 0 to 0.2 (P177L, Y305N, D322Y), from 0.2 to 0.6 (V92A, G349E, Y391C) and >0.6 (N36T, Q57E, A250V, N283D), with P177L and Y305N showing the lowest values (Figure 5B, upper panel). With the exclusion of N36 and Q57 positions, respectively absent in the mature protein or not resolved in the 6MV4 structure, analysis on protein stability, performed on each amino acid change combination (Supplementary Figure S3, lower panels), showed values ranging from −3 to −1 (P177L, Y305N, D322Y, Y391C), from −1 to 0 (V92A, A250V, G349E) and >0 (N283D), with Y305N showing the lowest value (Figure 5B, lower panel). The tolerated or damaging outputs, as well as the impact on protein stability, were in accordance, as expected, with the conservation degree of the original FIX residues (Supplementary Figure S4).

Noticeably, comparison between the calculated tolerance score and protein stability showed a good correlation (r = 0.63), with the Y305N and the N283D variants displaying extreme values (Figure 5C, left panel). The multitool prediction on natural variants (red circles), single- (orange circles) or nonsingle- (cyan circles) nucleotide changes, or on previously expressed variant (blue circle) of the catalytic S411 showed a calculated tolerance score equal to 0 for all substitutions, with the sole exception, albeit negligible, for arginine (0.066).

Interestingly, integration of prediction data based on the multitool approach allowed interpretation of inferred outputs, in terms of structure/functional properties, providing a classification of the 9 unclassified variants as severe (p.P177L, p.Y305N), severe/moderate (p.D322Y), moderate/mild (p.G349E, p.Y391C) and mild (p.N36T, p.Q57E, p.A250V, p.N283D), with the p.V92A variant classified as moderate/mild (Table 2).

4 | DISCUSSION

Dissection of genotype–phenotype relationships, and thus the understanding of the effects of nucleotide/amino acid changes,

corresponding levels shown as % of FIX-WT (right panel). (D) Western blotting analysis of FIX-WT and FIX-V92A protein forms in medium (M) and cell lysates (L) with anti-FIX antibodies (upper panels), as well as on cell lysates, normalized for total protein content, with anti-BiP and anti-GAPDH (loading control) antibodies (lower panels). The BiP/GAPDH ratio was obtained by densitometric analysis (Image Laboratory Software v4.0) of protein bands. (E) Coagulation times measured by APTT-based assays in FIX-depleted plasma (left panel), and evaluation of activity levels and specific activity (referred as the ratio between activity and Ag) of FIX-WT and FIX-V92A variants (right panel). (F) Western blotting analysis with detection of protein forms by anti-FIX antibodies (upper panel, with incubation time points indicated on top) obtained upon time-course incubation of FIX variants with plasma-derived FXIa. Activation profile over time is shown as relative band intensity, obtained by the ratio between densitometric value of each band on that of the last time point (lower panel). (G) Western blotting analysis with detection of protein forms by anti-FX antibodies (upper panel, with incubation time points indicated on top) obtained in time-course setting of the commercial chromogenic assay. Activation profile over time is shown as band intensity (middle panel) and relative band intensity (inset) or chromogenic activity as OD detection (lower panel) and relative activity (inset). Relative band intensity and activity were calculated on the last time point as described in (F). Results in (C) and (E) are reported as mean ± SD (n = 5 replicates). $^{**}P < .01$; $^{***}P < .001$; $^{****}P < .0001$.

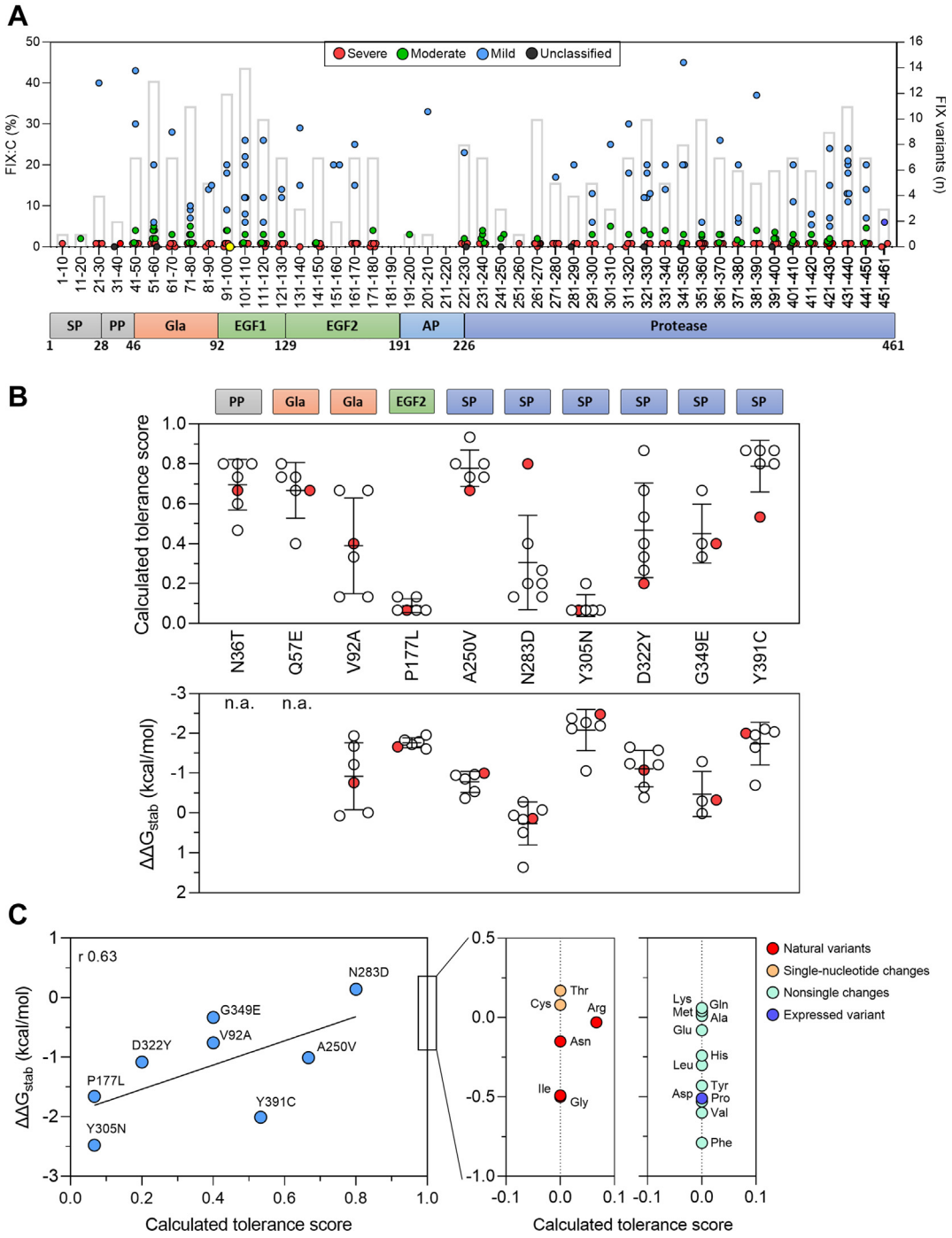

FIGURE 5 Database inspection and multitool analysis on unclassified FIX missense variants. (A) Distribution of factor (F)IX activity levels (FIX coagulant; colored circles, left y-axis) and number (n; grey bars, right y-axis) of reported unique (n = 1 patient) FIX missense variants (EAHAD, CHBMP, and HGMD databases) grouped as intervals of 10 amino acids in the FIX sequence (schematic representation of FIX domain organization and delimiting residues depicted below). HB severity is indicated as red (severe), green (moderate), or blue (mild) circles. Unclassified variants, including the p.V92A (yellow circle), are indicated by black circles. SP, signal peptide; PP, pro-peptide; EGF2, epidermal growth factor-like domain 2; AP, activation peptide. (B) Prediction of impact of amino acid substitutions, associated with HB (red circles) or deriving from single-nucleotide changes (white circles), by multitool (upper panel) and protein stability (lower panel) analysis of selected unclassified FIX missense variants with n = 1 or no overlapping amino acid substitutions (see Supplementary Table S2). Error bars indicate mean ± SD of calculated tolerance score (obtained as described in Figure 4B) or $\Delta\Delta G_{stab}$ (kcal/mol) values. The relative domain localization of each FIX position is schematically indicated on top; n.a., not available. (C) Correlation analysis between calculated tolerance score and protein stability for selected unclassified FIX missense variants (left panel). The rectangle on the right of the graph is referred to the interval (ranging from −1 to 0.5) of $\Delta\Delta G_{stab}$ values predicted for amino acid substitutions associated with HB (red circles) and derived from single- (orange circles) or nonsingle- (cyan circles) nucleotide changes, as well as of a previously expressed recombinant variant [17] (blue circle), at position 411 (serine of the catalytic triad, c195) (right panel).

TABLE 2 Proposed classification of unclassified factor (F)IX missense variants by integration of data from multitool analysis.

Genetic data[a]				Bioinformatic prediction		Inferred output		
Unclassified variant	Nucleotide change	Exon	Domain	Stability ($\Delta\Delta G_{stab}$)	Multitool (calculated tolerance score)	Secretion[b]	Function[c]	Severity
p.N36T	c.107A>C	2	Prepro	n.a.	0.67	n.a.	+++	Mild
p.Q57E	c.169C>G	2	Gla	n.a.	0.67	n.a.	+++	Mild
p.V92A	c.275T>C	3	Gla	−0.76	0.4	+	++	Moderate/mild
p.P177L	c.530C>T	6	Protease	−1.66	0.067	+	-	Severe
p.A250V	c.749C>T	7	Protease	−1.01	0.67	+	+++	Mild
p.N283D	c.847A>G	8	Protease	0.14	0.8	+++	+++	Mild
p.Y305N	c.913T>A	8	Protease	−2.48	0.067	-	-	Severe
p.D322Y	c.964G>T	8	Protease	−1.08	0.2	+	+	Severe/ moderate
p.G349E	c.1046G>A	8	Protease	−0.33	0.4	++	++	Moderate/mild
p.Y391C	c.1172A>G	8	Protease	−2.00	0.53	+	++	Moderate/mild

n.a., not available.

[a] European Association of Hemophilia and Allied Disorders (EAHAD), CDC Hemophilia Project (CHBMP), and Human Genome Mutation (HGMD) databases.

[b] Impact on secretion, based on DDmut prediction, was defined as - ($\Delta\Delta G_{stab} < -2$), + (from −2 to −0.5), ++ (from −0.5 to 0), and +++ (>0).

[c] Impact on function, based on multitool analysis, was defined as - (<0.1), + (from 0.1 to 0.2), ++ (from 0.2 to 0.6), and +++ (>0.6).

represents both an issue and a goal in human genetic disorders. In the context of hemophilia, considering that i) missense variants represent the major class of reported mutations (48% in hemophilia A [HA], 55% in HB) corresponding to a high percentage of patients (∼60% in HA, ∼70% in HB), ii) the frequent *F8* IVS1 and IVS22 inversions, affecting overall 50% of severe HA patients [38–41], are not included in the main hemophilia (EAHAD; CHAMP, https://www.cdc.gov/ncbddd/hemophilia/champs.html) databases, which leads to overestimating the relative number of reported FVIII missense variants, and iii) the proportion of missense variants affecting the large FVIII B domain might be reported but are less associated with HA [42,43], suggests that deepening the knowledge of the structure/functional impact of missense variants is of particular relevance for HB. In this scenario, shedding light on challenging (ie, mild HB forms) or unclassified FIX missense variants would contribute elements to better understand HB phenotypic variability, as well as to depict a complete picture of HB-associated missense variants. Importantly, this wealth of knowledge, driven by classification of the underlying disease-associated defect, and thus of patients, has important implications for genetic counseling, diagnosis, management, and treatment.

Here, the identification of the unclassified p.V92A FIX variant affecting a FIX region at the interface between Gla and EGF1 domains, not resolved in FIX structures currently available, led us to challenge different modeling/prediction systems to provide a multiple-level analysis and characterization of this variant, which served as starting point to widen our approach to a panel of reported but unclassified FIX missense variants.

Interestingly, the progress of molecular modeling tools has been particularly notable in recent years owing to the introduction of coevolution-based contact prediction [12,14] and deep neural-network learning [15] techniques. Based on these approaches, the generated models of the p.V92A variant, associated with a functional defect (specific activity in patient plasma, 0.52) causing mild HB, suggested that the valine-to-alanine substitution alters the conformation of Gla and EGF1-EGF2 domains, affecting 2 structural requirements (Ca^{++} affinity and EGF1-2 conformation) being crucial for maintaining an optimal conformation of the serine protease domain and thus efficient FIXa catalytic activity [44]. In particular, the predicted perturbations, and particularly those affecting EGF1-2, were observed to reduce FIX interaction with FXIa, responsible for FIX activation [36,45,46], and to exert subtle rearrangements of FIXa catalytic triad. This is in line with previous studies showing that FIX EGF1-2 domains are involved in interactions with FXIa, as well as with the reporting of several mutations in EGF1 (>70) being associated with several forms of both type I and type II HB [19,47]. These impairments are in accordance with, and help interpret, the prolongation of coagulation times in APTT assays (APTT ratio >1.5) and the reduced thrombin generation capacity observed in patient plasma.

A multitool analysis and expression studies on the p.V92A variant showed i) a differential impact of amino acid substitutions arising from possible single-nucleotide changes, with significant correlation with K_{ow} values, excepting the aromatic phenylalanine, as well as with protein stability, and ii) reduced secreted protein and activity levels, with an estimated *in vitro* specific coagulant activity (∼0.5)

recapitulating that of patient variant, and data on intracellular protein (BiP/GAPDH ratio) suggesting a folding defect. Time-course results pointed toward affected FXIa-mediated activation as the major determinant underlying the molecular defect of the V92A variant, with a minor impact on the catalytic activity of FIX protease domain. These data support the predicted lower interaction with FXIa, and also provide evidence on the FXa activation capacity of the V92A variant, with insights into the efficiency of the intrinsic tenase complex, albeit indirectly. The challenged multitool approach indicated a moderate impact of the valine-to-alanine amino acid substitution at FIX position 92, whose output at the protein level can be inferred in terms of secretion (protein stability value) and function (calculated tolerance score). On the other hand, expression and functional studies provided experimental evidence contributing information to dissect the structural and functional properties of the p.V92A variant and, importantly, to validate the multitool analysis.

This prompted us to further challenge our approach by extending the analysis to unclassified FIX missense variants reported as unique (n patients = 1), as the p.V92A variant, and lacking information on FIX levels or HB phenotype. Among the 30 variants identified, for 20 of them we inferred the associated HB severity from that of overlapping missense variants. Paradigmatic examples are position R191, affecting one of the FIX activation sites and mainly associated with moderate/mild HB [17], position W453, highly conserved in FVII, FIX, FX, FXI, and protein C up to chymotrypsin and whose substitution/absence is associated with virtually undetectable protein levels [48–52], or amino acid changes affecting cysteine residues mainly impacting on protein stability. The selected FIX missense variants, and amino acid substitutions stemming from single-nucleotide changes, that were challenged with our integrated approach, showed a differential distribution of predicted values depending on relative position in FIX and conservation degree of the affected residue. The good correlation of calculated tolerance scores and protein stability identified 2 variants (Y305N and N283D) displaying extreme values, which suggested reference for severe and mild outputs helpful to interpret that of other variants. The Y305 position also provides support for the goodness of prediction, showing low values for those substitutions involving amino acids with relevant differences in biochemical properties (His, Asp, Ser, Cys), but with exception of phenylalanine ascribable to the shared aromatic side chain. To further validate the multitool prediction analysis, we challenged amino acid substitutions at the catalytic S411, namely i) mimicking natural FIX variants, associated with severe HB but compatible with protein secretion, ii) deriving from single- or nonsingle-nucleotide changes, and iii) producing the S411P variant, previously expressed and resulting in devoid-of-function but normally secreted FIX [17]. All substitutions of S411, as expected for a crucial residue for enzyme function, were predicted as detrimental in terms of functional output, as indicated by a calculated tolerance score equal to 0 excepting the negligible arginine value, and with a relatively moderate impact on protein stability (values ranging from −0.8 to 0.2). This suggested the definition of 0.1 as the lower limit in terms of functional properties, and the value of −1 as a reference for moderate impact in terms of protein stability. Overall, integration of predicted and evidence-based reference values resulted in inferred outputs on secretion and function that allowed us to propose a classification for the selected panel of unclassified FIX missense variants.

We are aware that prediction tools may have limitations. However, this mainly applies to single-tool predictions relying upon a single algorithm. With this in mind, we rationally took advantage of a panel of multiple tools (n = 15), and thus multiple algorithms, with resulting outputs being coherently defined to categorize the different reference intervals associated with each tool, which allowed us to obtain the calculated tolerance score. Importantly, correlation of this score with values obtained from the complementary prediction on protein stability, as well as the availability of evidence-based data (ie, from *in vitro* studies), provided important elements and reference values driving interpretation of the final protein output. Moreover, despite modeling of activation peptide residues is limited by the lack of available structures, and interaction with FXIa by docking with the isolated FXIa protease domain, which might lead to underestimating the contribution of exosite-mediated interactions, the newly generated FIX-V92A model supported the multitool analysis, as well as data on the characterization of the p.V92A variant in plasma and recombinant systems, thus providing novel molecular and structural insights into FIX domains/regions not resolved in currently available structures.

In conclusion, our data support the rational integration of prediction tool outputs, especially as multitool and multiparameter analyses, as a useful approach to contribute elements for the interpretation of genotype/phenotype as well as structure/function relationships of unclassified missense variants, also in other high-impact coagulation disorders (ie, HA), with implications for diagnosis, management, and treatment of hemophilia patients, and potentially translatable into other human disorders.

ACKNOWLEDGMENTS

Financial support from Fondazione Policlinico Universitario Gemelli IRCCS (Ricerca corrente-2023) (RDC) and Fondo di Ateneo per la Ricerca (FAR) (A.B., M.F.T.) is gratefully acknowledged.

AUTHOR CONTRIBUTIONS

M.S., M.F.T., A.F., M.T., M.B., and L.D.G. carried out the experiments and acquired the clinical data; M.F.T. created recombinant plasmids, performed expression studies as well as protein and functional assays; P.C. and A.M. performed the molecular diagnosis; C.S. and S.L. revised the manuscript; M.F.T. and A.B. performed database and multitool analysis; R.D.C. generated, analyzed and interpreted the molecular modeling data; R.D.C. and A.B. conceived the study and designed research, analyzed data and wrote the manuscript. All authors revised and approved the final version of the manuscript.

DECLARATION OF COMPETING INTERESTS

The authors declare not having any relationship with a person or organization that could affect their objectivity, or inappropriately influence their actions.

ORCID

Alessio Branchini https://orcid.org/0000-0002-6113-2694

REFERENCES

[1] Mannucci PM, Tuddenham EG. The hemophilias—from royal genes to gene therapy. *N Engl J Med*. 2001;344:1773–9.
[2] Bolton-Maggs PH, Pasi KJ. Haemophilias A and B. *Lancet*. 2003;361:1801–9.
[3] Goodeve AC. Hemophilia B: molecular pathogenesis and mutation analysis. *J Thromb Haemost*. 2015;13:1184–95.
[4] White GC 2nd, Rosendaal F, Aledort LM, Lusher JM, Rothschild C, Ingerslev J, Factor V, Factor IXS. Definitions in hemophilia. Recommendation of the scientific subcommittee on factor VIII and factor IX of the scientific and standardization committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost*. 2001;85:560.
[5] Venkateswaran L, Wilimas JA, Jones DJ, Nuss R. Mild hemophilia in children: prevalence, complications, and treatment. *J Pediatr Hematol Oncol*. 1998;20:32–5.
[6] Castaman G, Matino D. Hemophilia A and B: molecular and clinical similarities and differences. *Haematologica*. 2019;104:1702–9.
[7] Goodeve AC, Reitsma PH, McVey JH, Working Group on Nomenclature of the S, Standardisation Committee of the International Society on T. Haemostasis. Nomenclature of genetic variants in hemostasis. *J Thromb Haemost*. 2011;9:852–5.
[8] den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, Roux AF, Smith T, Antonarakis SE, Taschner PE. HGVS Recommendations for the description of sequence variants: 2016 update. *Hum Mutat*. 2016;37:564–9.
[9] Sacco M, Lancellotti S, Ferrarese M, Bernardi F, Pinotti M, Tardugno M, De Candia E, Di Gennaro L, Basso M, Giusti B, Papi M, Perini G, Castaman G, De Cristofaro R. Noncanonical type 2B von Willebrand disease associated with mutations in the VWF D'D3 and D4 domains. *Blood Adv*. 2020;4:3405–15.
[10] Sacco M, Lancellotti S, Branchini A, Tardugno M, Testa MF, Lunghi B, Bernardi F, Pinotti M, Giusti B, Castaman G, De Cristofaro R. The p.P1127S pathogenic variant lowers von Willebrand factor levels through higher affinity for the macrophagic scavenger receptor LRP1: clinical phenotype and pathogenic mechanisms. *J Thromb Haemost*. 2022;20:1818–29.
[11] Zhang J, Liang Y, Zhang Y. Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. *Structure*. 2011;19:1784–95.
[12] Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res*. 2005;33:2302–9.
[13] Lu CH, Chen CC, Yu CS, Liu YY, Liu JJ, Wei ST, Lin YF. MIB2: metal ion-binding site prediction and modeling server. *Bioinformatics*. 2022;38:4428–9.
[14] Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S. The ClusPro web server for protein-protein docking. *Nat Protoc*. 2017;12:255–78.
[15] Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci U S A*. 1992;89:2195–9.
[16] Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*. 2016;32:3676–8.
[17] Branchini A, Morfini M, Lunghi B, Belvini D, Radossi P, Bury L, Serino ML, Giordano P, Cultrera D, Molinari AC, Napolitano M, Bigagli E, Castaman G, Pinotti M, Bernardi F, Ge PSGoA. F9 missense mutations impairing factor IX activation are associated with pleiotropic plasma phenotypes. *J Thromb Haemost*. 2022;20:69–81.
[18] Ferrarese M, Pignani S, Lombardi S, Balestra D, Bernardi F, Pinotti M, Branchini A. The carboxyl-terminal region of human coagulation factor X as a natural linker for fusion strategies. *Thromb Res*. 2019;173:4–11.
[19] Rallapalli PM, Kemball-Cook G, Tuddenham EG, Gomez K, Perkins SJ. An interactive mutation database for human coagulation factor IX provides novel insights into the phenotypes and genetics of hemophilia B. *J Thromb Haemost*. 2013;11:1329–40.
[20] McVey JH, Rallapalli PM, Kemball-Cook G, Hampshire DJ, Giansily-Blaizot M, Gomez K, Perkins SJ, Ludlam CA. The European Association for Haemophilia and Allied Disorders (EAHAD) Coagulation Factor Variant Databases: important resources for haemostasis clinicians and researchers. *Haemophilia*. 2020;26:306–13.
[21] Li T, Miller CH, Payne AB, Craig Hooper W. The CDC Hemophilia B mutation project mutation list: a new online resource. *Mol Genet Genomic Med*. 2013;1:238–45. https://doi.org/10.1002/mgg3.30
[22] Krawczak M, Ball EV, Fenton I, Stenson PD, Abeysinghe S, Thomas N, Cooper DN. Human gene mutation database-a biomedical information and research resource. *Hum Mutat*. 2000;15:45–51.
[23] Zhou Y, Pan Q, Pires DEV, Rodrigues CHM, Ascher DB. DDMut: predicting effects of mutations on protein stability using deep learning. *Nucleic Acids Res*. 2023;51:W122–8.
[24] Weigt M, White RA, Szurmant H, Hoch JA, Hwa T. Identification of direct residue contacts in protein-protein interaction by message passing. *Proc Natl Acad Sci U S A*. 2009;106:67–72.
[25] Marks DS, Colwell LJ, Sheridan R, Hopf TA, Pagnani A, Zecchina R, Sander C. Protein 3D structure computed from evolutionary sequence variation. *PLoS One*. 2011;6:e28766. https://doi.org/10.1371/journal.pone.0028766
[26] Mortuza SM, Zheng W, Zhang C, Li Y, Pearce R, Zhang Y. Improving fragment-based ab initio protein structure assembly using low-accuracy contact-map predictions. *Nat Commun*. 2021;12:5011.
[27] Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, Zidek A, Nelson AWR, Bridgland A, Penedones H, Petersen S, Simonyan K, Crossan S, Kohli P, Jones DT, Silver D, Kavukcuoglu K, Hassabis D. Improved protein structure prediction using potentials from deep learning. *Nature*. 2020;577:706–10.
[28] Baron M, Norman DG, Harvey TS, Handford PA, Mayhew M, Tse AG, Brownlee GG, Campbell ID. The three-dimensional structure of the first EGF-like module of human factor IX: comparison with EGF and TGF-alpha. *Protein Sci*. 1992;1:81–90.
[29] Vadivel K, Schreuder HA, Liesum A, Schmidt AE, Goldsmith G, Bajaj SP. Sodium-site in serine protease domain of human coagulation factor IXa: evidence from the crystal structure and molecular dynamics simulations study. *J Thromb Haemost*. 2019;17:574–84.
[30] Radisky ES, Lee JM, Lu CJ, Koshland DE Jr. Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates. *Proc Natl Acad Sci U S A*. 2006;103:6835–40.
[31] Davis LM, McGraw RA, Ware JL, Roberts HR, Stafford DW. Factor IXAlabama: a point mutation in a clotting protein results in hemophilia B. *Blood*. 1987;69:140–3.
[32] Lin SW, Smith KJ, Welsch D, Stafford DW. Expression and characterization of human factor IX and factor IX-factor X chimeras in mouse C127 cells. *J Biol Chem*. 1990;265:144q50.
[33] Mathur A, Bajaj SP. Protease and EGF1 domains of factor IXa play distinct roles in binding to factor VIIIa. Importance of helix 330 (helix 162 in chymotrypsin) of protease domain of factor IXa in its interaction with factor VIIIa. *J Biol Chem*. 1999;274:18477–86.
[34] Knobe KE, Persson KE, Sjorin E, Villoutreix BO, Ljung RC. Functional analysis of the factor IX epidermal growth factor-like domain mutation Ile66Thr associated with mild hemophilia B. *Pathophysiol Haemost Thromb*. 2006;35:370–5.
[35] Lenting PJ, Christophe OD, Maat H, Rees DJ, Mertens K. Ca2+ binding to the first epidermal growth factor-like domain of human blood coagulation factor IX promotes enzyme activity and factor VIII light chain binding. *J Biol Chem*. 1996;271:25332–7.

[36] Brandstetter H, Bauer M, Huber R, Lollar P, Bode W. X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and hemophilia B. *Proc Natl Acad Sci U S A*. 1995;92:9796–800.
[37] Lai CW, Aronson DE, Snapp EL. BiP availability distinguishes states of homeostasis and stress in the endoplasmic reticulum of living cells. *Mol Biol Cell*. 2010;21:1909–21.
[38] Cumming AM, Network UKHCDOHGL. The factor VIII gene intron 1 inversion mutation: prevalence in severe hemophilia A patients in the UK. *J Thromb Haemost*. 2004;2:205–6.
[39] Schroder J, El-Maarri O, Schwaab R, Muller CR, Oldenburg J. Factor VIII intron-1 inversion: frequency and inhibitor prevalence. *J Thromb Haemost*. 2006;4:1141–3.
[40] Bagnall RD, Giannelli F, Green PM. Int22h-related inversions causing hemophilia A: a novel insight into their origin and a new more discriminant PCR test for their detection. *J Thromb Haemost*. 2006;4:591–8.
[41] De Brasi CD, Bowen DJ. Molecular characteristics of the intron 22 homologs of the coagulation factor VIII gene: an update. *J Thromb Haemost*. 2008;6:1822–4.
[42] Ogata K, Selvaraj SR, Miao HZ, Pipe SW. Most factor VIII B domain missense mutations are unlikely to be causative mutations for severe hemophilia A: implications for genotyping. *J Thromb Haemost*. 2011;9:1183–90.
[43] Testa MF, Lombardi S, Bernardi F, Ferrarese M, Belvini D, Radossi P, Castaman G, Pinotti M, Branchini A. Translational readthrough at F8 nonsense variants in the factor VIII B domain contributes to residual expression and lowers inhibitor association. *Haematologica*. 2023;108:472–82.
[44] Schmidt AE, Bajaj SP. Structure-function relationships in factor IX and factor IXa. *Trends Cardiovasc Med*. 2003;13:39–45.
[45] Gailani D, Geng Y, Verhamme I, Sun MF, Bajaj SP, Messer A, Emsley J. The mechanism underlying activation of factor IX by factor XIa. *Thromb Res*. 2014;133(Suppl 1):S48–51.
[46] Ngo JC, Huang M, Roth DA, Furie BC, Furie B. Crystal structure of human factor VIII: implications for the formation of the factor IXa-factor VIIIa complex. *Structure*. 2008;16:597–606.
[47] Giannelli F, Green PM, Sommer SS, Poon M, Ludwig M, Schwaab R, Reitsma PH, Goossens M, Yoshioka A, Figueiredo MS, Brownlee GG. Haemophilia B: database of point mutations and short additions and deletions—eighth edition. *Nucleic Acids Res*. 1998;26:265–8.
[48] Kurachi S, Pantazatos DP, Kurachi K. The carboxyl-terminal region of factor IX is essential for its secretion. *Biochemistry*. 1997;36:4337–44.
[49] Katsumi A, Kojima T, Senda T, Yamazaki T, Tsukamoto H, Sugiura I, Kobayashi S, Miyata T, Umeyama H, Saito H. The carboxyl-terminal region of protein C is essential for its secretion. *Blood*. 1998;91:3784–91.
[50] Branchini A, Baroni M, Burini F, Puzzo F, Nicolosi F, Mari R, Gemmati D, Bernardi F, Pinotti M. The carboxyl-terminal region is NOT essential for secreted and functional levels of coagulation factor X. *J Thromb Haemost*. 2015;13:1468–74.
[51] Ferrarese M, Baroni M, Della Valle P, Spiga I, Poloniato A, D'Angelo A, Pinotti M, Bernardi F, Branchini A. Missense changes in the catalytic domain of coagulation factor X account for minimal function preventing a perinatal lethal condition. *Haemophilia*. 2019;25:685–92.
[52] Branchini A. The carboxyl-terminal region of coagulation serine proteases: a matter of cut and change. *J Thromb Haemost*. 2021;19:917–9.

SUPPLEMENTARY MATERIAL

The online version contains supplementary material available at https://doi.org/10.1016/j.jtha.2024.07.008