Human Genomics

RESEARCH Open Access

Molecular mechanisms and therapeutic strategies for the recurrent *F9* (c.520 + 13 A > G) variant in hemophilia B

Huayang Zhang[1], Chong Wang[1], Meixiu Gu[1], Zhimin Meng[1], Yichao Guo[1], Weitao Zhang[1], Dario Balestra[2*], Wei Guo[1,3,4,5*] and Beili Wang[1,3,4,5*]

Abstract

Background Hemophilia B (HB), an X-linked recessive disorder, results from variants in the coagulation factor IX gene (*F9*). The *F9* c.520 + 13 A > G variant is a recurrent intronic variant in HB patients, accounting for 15.05% of all documented *F9* intronic variants. Despite prior predictions of its impact, the molecular mechanism associated with moderate to mild HB remains undissected.

Materials and methods In silico predictions and splicing-competent cDNA constructs were used to assess the impact of *F9* c.520 + 13 A > G on mRNA splicing. Factor IX (FIX) variant (p.V174delinsGHNLM) expression in HEK293T cells was evaluated using activated partial thromboplastin time, enzyme-linked immunosorbent assay, Western blot analysis, and immunofluorescence analyses. Structural modeling and molecular dynamics simulations were performed to evaluate the structural impact of the variant. Engineered U1 small nuclear RNA (U1snRNA) was challenged with *F9* full-length splicing-competent constructs to evaluate splicing correction.

Results The *F9* c.520 + 13 A > G variant caused nearly complete aberrant splicing, producing the *F9* c.520_521insGTCATAATCTGA insertion and the in-frame FIX p.V174delinsGHNLM variant. A small amount (approximately 10%) of wild-type FIX was also detected. We characterize the p.V174delinsGHNLM variant, which exhibited impaired secretion and increased intracellular accumulation. Interestingly, an engineered U1snRNA partially rescued aberrant splicing, restoring functional FIX levels to approximately 40%.

Conclusion This study elucidates the molecular mechanism of the *F9* c.520 + 13 A > G variant, which activates a cryptic 5′ splice site in intron 5, leading to an in-frame FIX (p.V174delinsGHNLM) with secretion defects and loss of protein function. And Engineered U1snRNA partially rescued the splicing defect.

*Correspondence:
Dario Balestra
blsdra@unife.it
Wei Guo
guo.wei@zs-hospital.sh.cn
Beili Wang
wang.beili1@zs-hospital.sh.cn

Full list of author information is available at the end of the article

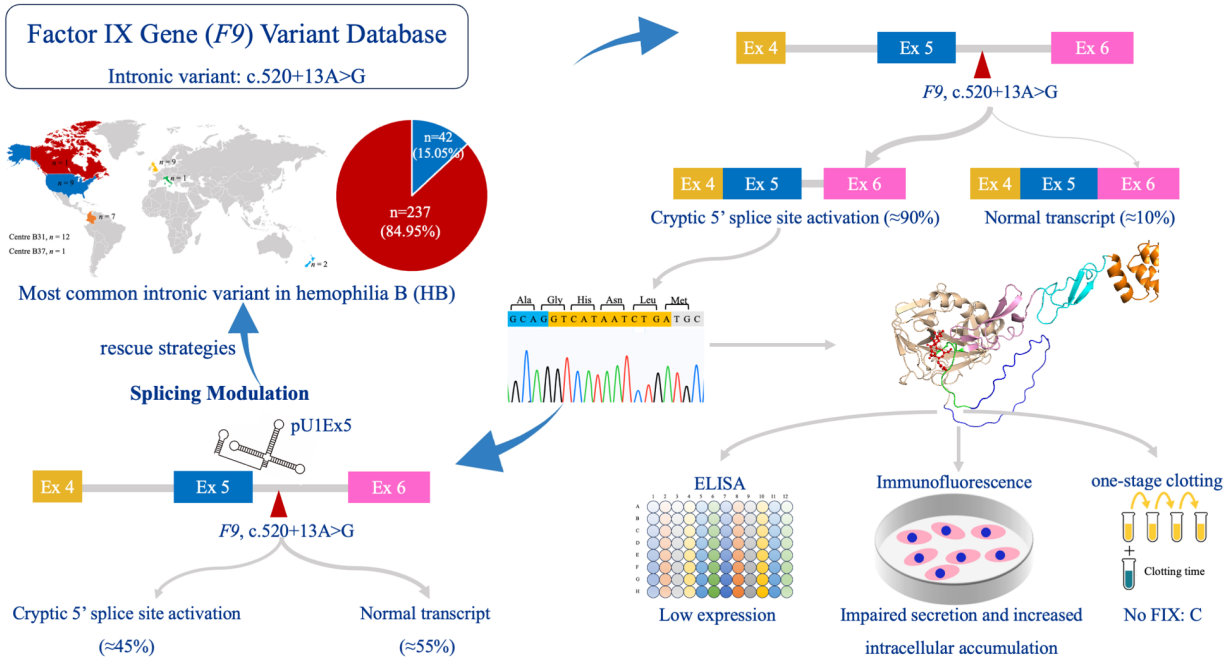
Highlights

- **What is known on this topic?**1.The *F9* c.520 + 13 A > G variant is a recurrent intronic variant associated with hemophilia B.
- 2.The molecular mechanism has been only hypothesized, but never experimentally investigated.

- **What does this paper add?**1.The *F9* c.520 + 13 A > G variant induces aberrant splicing through activation of an intronic cryptic 5′ splice site.
- 2.The insertion produces an in-frame FIX (p.V174delinsGHNLM) with impaired secretion, increased intracellular accumulation, and reduced coagulation activity.
- 3.Engineered U1 small nuclear RNA (U1snRNA) successfully rescues the splicing defect, providing a tailored RNA-based therapeutic strategy.

Keywords Hemophilia B, Intronic variant, Pathogenic mechanism, Aberrant pre-mRNA splicing, U1snRNA

Graphical abstract



Visual Summary. The *F9* c.520+13A>G variant causes aberrant splicing via a cryptic 5′ splice site, leading to a dysfunctional FIX protein with secretion defects. Engineered U1snRNA successfully corrects the splicing defect, restoring functional FIX expression.

Introduction

Hemophilia B (HB) is a hereditary X-linked recessive bleeding disorder with an incidence of 1 in 30,000 male births [1]. HB is caused by a deficiency in coagulation factor IX (FIX), either quantitative or qualitative. Based on residual FIX activity (FIX: C) levels in the plasma, the HB phenotype is classified as mild (>5–40 IU/dL), moderate (1–5 IU/dL), or severe (< 1 IU/dL) [2]. FIX is encoded by the *F9* gene, located on chromosome Xq27.1 and spanning 32.7 kb. *F9* comprises 8 exons and encodes a 2.8-kb transcript [3].

F9 variants contribute to FIX deficiency through multiple mechanisms, including alterations in gene structure [4], splicing [5], translation [6], post-translational modifications [7], protein folding [8], and formation of a functional complex [9]. The FIX gene (*F9*) variant database (https://www.factorix.org/) catalogs variants from over 5000 unrelated patients, and includes 1692 distinct gene changes [10]. Intronic point variants typically lead to aberrant splicing, resulting in alteration of protein expression [11]. According to the *F9* variant database, many HB patients with intronic variants exhibit severe to moderate phenotypes, consistent with this mechanism [12]. Currently, the *F9* variant database reports 279 patients with intronic point variants, accounting for 6.08% of all patients [10]. Among these is the recurrent *F9* intronic variant c.520 + 13 A >G (15.05% of the total), which is associated with moderate to mild HB phenotypes [13–16].

P.M. Green et al. [17] conducted a systematic study on HB variants in the United Kingdom (UK) population, demonstrating mutational events that strongly support the deleterious nature of the *F9* c.520 + 13 A >G variant. Although splice site alterations and their potential impact on FIX structure have been predicted, further experimental validation and clinical studies are required to fully determine the physiological and clinical implications of this variant.

The high prevalence of c.520 + 13 A >G and its predicted molecular mechanism make it a suitable target for RNA therapeutics, like U1 small nuclear RNA (U1snRNA), and in some contexts, antisense U7 small nuclear RNA (U7snRNA). U1snRNA, the RNA component of the spliceosomal U1 ribonucleoprotein (U1snRNP), mediates recognition of the 5′ splice site (5′ss) and exon definition during the initial splicing step [18]. It has emerged as a promising therapeutic for human diseases caused by splicing variants, which are often associated with severe forms in over 15% of cases [19]. U1snRNA variants with increased complementarity to the 5′ss of the defective exon (compensatory U1snRNA) or those targeting downstream intronic sequences (exon-specific U1snRNA, ExSpeU1) can rescue mRNA splicing. U7 snRNA, typically involved in histone pre-mRNA 3′-end processing, can be converted into a versatile tool for splicing modulation with a small change in the binding site for Sm/Lsm proteins [20, 21]. When embedded into an snRNP particle, the antisense sequence is protected from degradation and accumulates in the nucleus, where splicing occurs. U1snRNA has consistently shown strong potential in rescuing splicing in various cellular models of human genetic disease, including coagulation factor deficiencies [18, 19]. U7snRNA can also be engineered for splicing modulation, although its efficiency may be more context-dependent [22].

This study employed in silico prediction and splicing-competent FIX expression assays to investigate the molecular mechanisms of the *F9* c.520 + 13 A > G. This variant activates a cryptic 5′ splice site in intron 5, resulting in aberrant splicing and the in-frame FIX p.V174delinsGHNLM. Characterization of the FIX p.V174delinsGHNLM showed impaired secretion and increased intracellular accumulation, accompanied by markedly reduced coagulant activity. Furthermore, in a splicing-competent FIX expression system, an engineered U1snRNA-based correction strategy successfully rescued splicing and protein activity. This correction approach lays a strong foundation for future gene therapy strategies aimed at treating HB.

Materials and methods

Variant nomenclature

Variants and alternative transcripts were annotated following the guidelines of the Human Genome Variation Society (HGVS) (http://hgvs-nomenclature.org), based on the *F9* reference sequences (RefSeq NM_000133.4 and NP_000124.1). Sequence variant descriptions were verified using the Mutalyzer 3 online tool (https://mutalyzer.nl/; accessed on October 28, 2025).

Bioinformatic tools

Human Splicing Finder (HSF) [23], MaxEntScan (MES) [24], NNSplice [25], and SpliceAI [26] were employed to predict the effects of variants on splicing.

HSF employs "position weight matrices" constructed from aligned consensus sequences to weigh each nucleotide in a sequence and uses the sum of individual nucleotide's weights to derive its potential splice site strength [23]. MES scores sequences for splice sites by employing the "maximum entropy principle" that takes into account adjacent and nonadjacent position dependencies in consensus splice sites [24]. NNSPLICE employs two "neural networks" that were trained on consensus splice sites, while also considering dinucleotide frequencies due to the strong correlation between neighboring nucleotides in splice site consensus sequences [25]. SpliceAI, a deep residual neural network, learns splicing determinants directly from the primary sequence by evaluating 10,000 nucleotides of the flanking context sequence to predict the splice function of each position in the pre-mRNA transcript [26].

Splicing and expression vectors construction

Expression vectors for the wild-type splicing-competent hFIX cDNA cassette (pBSK-FIX-WT) were previously reported [27]. Briefly, the pBSK-FIX-WT construct was built on the pCDNA 3.1 + backbone, using a simian virus 40 (SV40) promoter to drive expression. The cassette includes the human FIX cDNA sequence covering exons 1–4 with 544 bp of intron 4, followed by an *Nde* I restriction site, then the last 314 bp of intron 4, exon 5, and 278 bp of intron 5. Another *Nde* I site precedes the 908 bp segment of intron 5, followed by cDNA for exons 6–8, ending with the SV40 polyadenylation signal. For generating variants to investigate rescue at the protein level, the *Nde* I - *Nde* I fragment containing the last part of intron 4, exon 5, and a portion of intron 5 was subcloned into the FIX cDNA sequences. In addition to the splicing-competent pBSK-FIX-WT, we constructed the optimized pBSK-FIX (pBSK-FIX-OPT-WT), incorporating variants at the acceptor splice site to ensure exon inclusion in wild-type conditions, allowing for assessment of alternative splicing. The *F9* c.520 + 13 A >G variant was introduced using site-directed mutagenesis.

The modified pU1snRNA (pU1Ex5) and pU7snRNA (pU7c.520 + 13G) variant plasmids were generated by homologous recombination with targeted oligonucleotides (BGI, Shenzhen, China), based on the previously described U1snRNA [28] and U7snRNA vectors [22].

For recombinant FIX expression, the wild-type human FIX cDNA was synthesized and cloned into the pcDNA3.1(+) vector to construct the wild-type expression plasmid (FIX-WT). Recombinant FIX p.V174delinsGHNLM expression vectors were subsequently generated from FIX-WT by homologous recombination using synthetic DNA fragments carrying the desired mutations. All constructs were confirmed by DNA sequencing.

In vitro splicing assays

Human embryonic kidney 293T (HEK293T) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco, Grand Island, NY, USA) supplemented with 10% fetal bovine serum (FBS) (Gibco) at 37 ℃ with 5% CO_2. Cells were transiently transfected with splicing-competent *F9* constructs (pBSK-FIX-WT, pBSK-FIX-c.520 + 13 A >G, pBsK-FIX-opt-WT, or pBsK-FIX-opt-c.520 + 13 A >G) with or without pU1snRNA using linear polyethylenimine (PEI, MW 4000; Yeasen, China). Total RNA was extracted from the cells using TRIzol Reagent (Invitrogen), followed by phenol-chloroform extraction after 24 h of transfection. For reverse transcription polymerase chain reaction (RT-PCR), 2 μg of RNA was processed using the HiScript II First-Strand cDNA Synthesis Kit (Vazyme, Nanjing, China). PCR amplification utilized primers located on exons 4 and 6 of the pBSK-FIX constructs, with products separated on a 2% agarose gel. Cloning was performed with the pMD19-T Simple Vector (Takara, Dalian, China), and twenty clones per variant were sequenced. Denaturing capillary electrophoresis and fluorescent PCR were used to evaluate the relative ratios of multiple transcripts [29].

In vitro expression assays

The expression vectors include FIX-WT, FIX p.V174delinsGHNLM, pBSK-FIX-OPT-WT, and pBSK-FIX-OPT variant constructs, which were transiently transfected into HEK293T cells using linear polyethylenimine (PEI, MW 4000; Yeasen, China) following the manufacturer's instructions. After 24 h, the culture media were replaced with fresh DMEM (Gibco) with 1 × Insulin Transferrin Selenium Supplement (Beyotime), and Vitamin K1 (5 μg/mL). The media were harvested after 48 h and centrifuged for 5 min at 3000 × g. The levels of secreted FIX antigen (FIX: Ag) and FIX activity (FIX: C) were assessed using enzyme-linked immunosorbent assays (ELISA) and one-stage clotting method based on activated partial thromboplastin time (aPTT) according to the manufacturer's instructions.

Western blot analysis

The conditioned medium was collected, and cell debris was removed. Transfected cells were lysed using RIPA cell lysis buffer (Cell Signaling Technology, Beverly, MA, USA) at 4 °C, and the resulting lysates were stored at − 80 °C for subsequent analysis. Both the collected culture media and cell lysates from transfected HEK293T cells were mixed with 1 × SDS loading buffer and then heated at 95 °C for 5 min to denature the protein molecules. The prepared samples were then separated using SDS-PAGE and transferred onto the nitrocellulose membrane (Merck Millipore, Billerica, MA, USA). To block non-specific binding, the membranes were incubated in 5% non-fat milk (Sangon Biotech, Shanghai, China) for 60 min at room temperature. Following this, the membranes were exposed to primary polyclonal sheep antibody against human FIX (SAFIX-IG, Affinity Biologicals, Ancaster, ON, Canada) or GAPDH (Servicebio, China) at 1:2000 dilution overnight at 4 ℃. Subsequently, the membranes were incubated with horseradish peroxidase-conjugated secondary antibodies at a 1:2000 dilution for 60 min at room temperature. The blots were visualized using Enhanced Chemiluminescence Western Blotting Detection Reagents (Invitrogen), and the images were detected using the Tanon-5500 Imaging System (Tanon, Shanghai, China). GAPDH was used as a loading control in parallel blots to ensure equal protein loading.

Immunofluorescence staining of HEK293T cells expressing FIX

HEK239T cells were seeded onto poly-lysine-coated glass coverslips and transfected with either FIX WT or FIX p.V174delinsGHNLM expression vectors. Following 24 hours of transfection, the transfected cells were fixed with 4% formaldehyde at room temperature for 30 minutes. Subsequently, the fixed cells were incubated with the blocking buffer (1 × PBS, 5% goat serum, 0.3% Triton™ X-100) for 60 minutes to prevent non-specific binding. The cells were then exposed to a polyclonal sheep anti-human FIX antibody (Affinity Biologicals) at a dilution of 1:5000. Simultaneously, the cells were incubated with Calreticulin (D3E6) XP® Rabbit mAb (Cell Signaling Technology) at a dilution of 1:400 to stain the endoplasmic reticulum (ER), this incubation was at 4°C overnight. After the primary antibody incubation, the cells were incubated with Donkey Anti-Sheep IgG H&L (Alexa Fluor® 488) (Abcam) to visualize the expressed FIX and its intracellular localization. Additionally, the Anti-rabbit IgG (H + L), F (ab')2 Fragment (Alexa Fluor® 555 Conjugate) at a dilution of 1:500 was utilized to visualize the ER. For visualizing the cell nuclei,

4′,6-diamidino-2-phenylindole (DAPI) staining was employed. Finally, the coverslips were mounted under a Leica STELLARIS 5 high-resolution white-light confocal microscope (Leica Microsystems, Germany).

Molecular dynamics simulations

Models of full-size and FIX p.174delinsGHNLM were generated using AlphaFold 3 [30]. Five models were generated with predicted local distance difference test (pLDDT) >70, expected to be modeled with good accuracy. All models shown were rendered with the PyMOL software (Version 2.1, https://www.pymol.org/).

The selection of simulation parameters is crucial in accurately modeling molecular dynamics. We chose a static temperature of 300 K and standard pressure (1 Bar) for our simulation environment, as these conditions closely approximate the physiological environment, enabling a realistic simulation of molecular behavior. For the force field, we utilized Amber99sb-ildn, which has been widely validated and is regarded for its accuracy in describing biomolecular dynamics. To simulate the solvent environment, we applied the Tip3p water model and added an appropriate number of Na^+ ions to neutralize the system's total charge, ensuring stability and accuracy in the simulation.

The simulation process began with an energy minimization step using the steepest descent method to eliminate potential conformational stress and any unrealistic geometric configurations. This was followed by equilibration in both the isothermal-isochoric (NVT) ensemble and isothermal-isobaric (NPT) ensemble for 100,000 steps each, with coupling constants of 0.1 ps for a duration of 100 ps, to stabilize the system's temperature and pressure. Finally, we conducted a production molecular dynamics simulation with a total of 50,000,000 steps, using a 2 fs time step over a total simulation time of 100 nanoseconds to thoroughly capture the dynamic behavior of the system.

To analyze the simulation results, we calculated several key physical quantities, including Root Mean Square Deviation (RMSD) of the complex, Root Mean Square Fluctuation (RMSF), and Radius of Gyration (Rg).

Clinical classification of *F9* c.520 + 13 A > G variant

The clinical classification of the *F9* c.520 + 13 A > G variant was performed following the ClinGen Coagulation Factor Deficiency VCEP guidelines for *F9* (version 1.0.0, released 10/5/2023).

Statistical analysis

Abnormal splicing and protein expression percentages were calculated, with error bars representing the standard error of the mean (SEM, $n = 3$).

Results

Intronic variants account for over 6% of the reported *F9* alterations listed in FIX variant databases [10, 31]. These variants are associated with heterogeneous phenotypes, with severe forms occurring when FIX coagulant levels fall below 1% of normal. Among these, we focused on the *F9* c.520 + 13 A >G variant, the most frequent intronic variant ($n = 42$), which is associated with heterogeneous HB phenotypes ranging from mild to severe (Fig. 1, and Table S1).

Bioinformatics analysis on splicing

The single-base substitution c.520 + 13 A >G in *F9* intron 5 has been previously predicted as a splicing variant [17]. To investigate the impact of the c.520 + 13 A >G variant on *F9* mRNA splicing, we used several in silico tools, predicting the impact on splicing in complementary ways. We investigated the impact of the c.520 + 13 A >G variant on *F9* mRNA using in silico analysis (Table 1). In the presence of the variant, SpliceAI predicted a strong cryptic 5′ss (score 0.92) located 12 nucleotides downstream of the authentic 5′ss, suggesting this variant might alter the normal mRNA splicing process. Additionally, HSF and MES prediction tools indicated moderate to severe disruptions of the normal mRNA splicing process, whereas NNSplice showed no effect. If the cryptic splice site is indeed activated, this new transcript, incorporating 12 additional nucleotides into the mature mRNA, would result in the in-frame FIX (p.V174delinsGHNLM). Overall, in silico prediction suggested that the c.520 + 13 A >G variant creates a cryptic 5′ss, which could lead to the synthesis of a longer in-frame FIX protein, potentially affecting protein structure and function.

Effects of c.520 + 13 A > G variant on *F9* mRNA splicing

We investigated the splicing impact of the *F9* c.520 + 13 A > G variant using splicing-competent FIX constructs (pBSK-FIX-WT and pBSK-FIX-c.520 + 13 A > G) transfected into HEK293T cells. In the original constructs, the 3′ splice site of exon 5 was relatively weak, resulting in low exon inclusion in WT and predominant aberrant splicing in the variant, producing a 12-nucleotide insertion at the 3′ end of exon 5 (Figure S1).

To accurately quantify splicing and assess rescue strategies, we generated optimized constructs with enhanced 3′ splice site strength (pBSK-FIX-OPT-WT and pBSK-FIX-OPT-c.520 + 13 A > G) (Fig. 2A). Only the 3′ss was modified, while the 5′ss and surrounding exonic sequences remained unchanged, preserving the authentic splicing context. The 3′ss was optimized because exon 5 is a weakly defined exon, and strengthening the 3′ss allows more robust evaluation of exon inclusion and the effects of U1snRNA rescue. HEK293T cells transfected

A

Variants (HGVS)	Location	Subjects' data			
		Subjects	FIX: C (%)	FIX: Ag (%)	Severity
c.520+13A>G	Intron 5	42	<1/1-20	4/5/6/10	Se/Mo/Mi

B

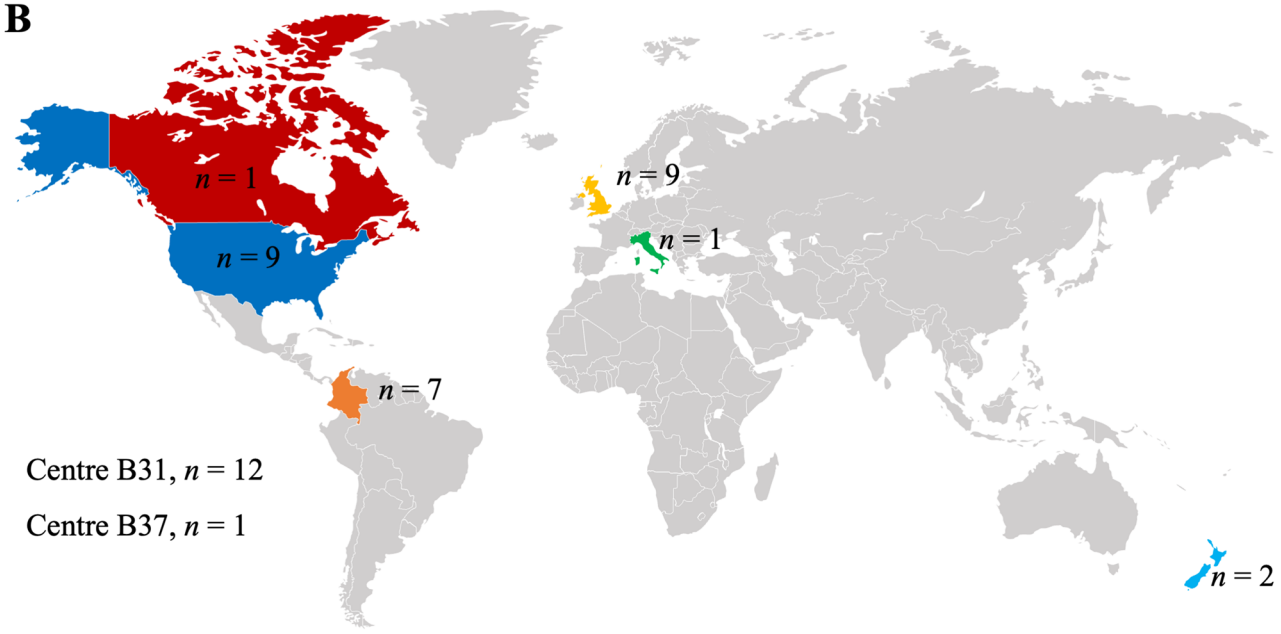


Fig. 1 Information on HB patients reported with the *F9* variant c.520 + 13 A > G. **A** The *F9* c.520 + 13 A > G variant, as reported in the FIX variant database (www.factorix.org). HGVS, Human Genome Variation Society; Se, severe; Mo, moderate; Mi, Mild. **B** Frequency distribution of the *F9* c.520 + 13 A > G variant. The number of patients from published reports is shown.

with these optimized constructs were analyzed 24 h post-transfection by RT-PCR using primers on exons 4 and 6. Electrophoresis revealed an additional band (~ 427 bp) in the variant, slightly larger than the expected 415 bp WT product (Fig. 2B). Sequencing confirmed the 12-nucleotide insertion (*F9* c.520_521insGTCATAATCTGA) resulting from activation of a cryptic 5′ splice site at c.520 + 12 (Fig. 2C).

Quantitative analysis using fluorescent-labeled RT-PCR amplicons and denaturing capillary electrophoresis showed that normal transcripts represented 8.96% of total products in the variant, whereas aberrant transcripts accounted for 91.04%. In contrast, the WT construct produced 100% normal splicing products (Fig. 2D). These results demonstrate that the c.520 + 13 A > G variant predominantly induces aberrant splicing, providing the basis for subsequent rescue experiments with engineered U1snRNA.

Recombinant FIX expression studies indicated a differential impact of amino acid changes on FIX protein

An aberrant transcript resulting from a 12-base insertion in intron 5 leads to an in-frame insertion variant (p.V174delinsGHNLM) in the FIX protein, specifically affecting the FIX linker region where amino acids are highly conserved adjacent to the insertion site (Fig. 3A). Quantitative analysis using ELISA revealed that the expression levels of the variant protein were markedly decreased (10.74 ± 1.76% of the FIX WT). This expression reduction was further substantiated by aPTT assays, which showed minimal detectable coagulation activity (0.83 ± 0.45% of the FIX WT) for the FIX p.V174delinsGHNLM variant (Fig. 3B).

Western blot analysis of the conditioned medium of HEK293T cells transfected with FIX WT revealed a band at ~ 66 kDa, corresponding to the full-length FIX protein (Fig. 3C). In contrast, the band was significantly reduced for conditioned media of the FIX p.V174delinsGHNLM variant. The densitometry analysis of the bands showed that the band intensity of FIX p.V174delinsGHNLM was

Table 1 In Silico prediction for splicing

Variants	Subjects' data				Splice sites prediction								
	No. of cases	FIX: C (%)	FIX: Ag (%)	Severity	Positions	HSF (0-100)		MES (0–12)		NNS (0–1)		SpliceAI (0-1)	
						A	V	A	V	A	V	DL	DG
c.520 + 13 A > G	42	< 1/1–20	4/5/6	Se/Mo/Mi	c.520[a]	81.44	81.44	3.00	3.00	0	0	0.00	0.92
					c.520 + 12[c]	52.50	79.64	0	5.11	0	0		

FIX: C FIX activity; *FIX: Ag* FIX antigen; *Mi* mild; *Mo* moderate; *Se* severe; a, the position of the authentic 5'ss; c, the position of the cryptic 5'ss. A, the scores of authentic splice sites. C, the scores of cryptic splice sites. MES, MaxEntScan, the score ranges from 0 to 12 for donor sites and from 0 to 16 for acceptor sites (threshold: ≥0); NNS, NNSplicer, the score ranges from 0 to 1; (threshold: ≥0.4); HSF, Human Splicing Finder, the score ranges from 0 to 100 (threshold: ≥65). Delta score predicted by SpliceAI, ranges from 0 to 1 (<0.35: weak predicted score; 0.35–0.8: intermediate predicted score; >0.8: high predicted score). DL donor loss; DG donor gain.

approximately 15% of the intensity of the FIX wild-type counterpart (Fig. 3D). Moreover, additional Western blot analyses of cell lysates from HEK293T cells transfected with either FIX WT or FIX p.V174delinsGHNLM indicated two bands between 52 and 66 kDa. Interestingly, the FIX p.V174delinsGHNLM showed a much stronger intracellular signal, approximately 2.5 times higher than FIX WT (Fig. 3C, D).

In the wild-type, rFIX-wt showed partial overlap with the ER marker, consistent with proper secretion and transport. In contrast, the FIX p.V174delinsGHNLM variant exhibited marked intracellular retention and increased co-localization with the ER marker (Figs. 3E). Western blot analysis further showed increased intracellular levels and reduced secretion of the variant compared with wild-type (Fig. 3C, D). These observations indicate that the variant protein has altered intracellular distribution and impaired secretion.

Molecular dynamics simulations of FIX

The primary sequence of human FIX consists of 461 residues (NP_000124.1). High-confidence structural models for full-length human FIX (NP_000124.1) and the p.V174delinsGHNLM variant were generated using AlphaFold3, with predicted TM-scores (pTM) of 0.75 for both WT and variant, indicating reliable modeling accuracy across most regions. Initial molecular models of the FIX WT and FIX p.V174delinsGHNLM were shown in Fig. 4A, B.

Both FIX WT and FIX p.V174delinsGHNLM systems equilibrated within the first 100 ns of the 200 ns simulations (Fig. 4C). In molecular dynamics simulations, the root mean square deviation (RMSD) serves as a key indicator of structural stability by quantifying the average displacement of atomic positions relative to the initial structure. During the simulation, FIX WT (red line) exhibited RMSD values around 10 Å, gradually reaching a stable plateau, indicating that the WT protein maintains a well-folded and stable conformation. In contrast, the FIX p.V174delinsGHNLM (blue line) showed lower RMSD values with pronounced fluctuations between 5 and 6 Å, suggesting that the insertion induces transient structural deviations and reduced overall stability compared with the WT. Analysis of the radius of gyration (Rg), which reflects the compactness of the protein structure, revealed clear differences between the two systems (Fig. 4D). FIX WT gradually decreased in Rg during the initial phase of the simulation, stabilizing at 27–30 Å after 125 ns, consistent with a compact and stable fold. In contrast, the FIX p.V174delinsGHNLM maintained a higher and more variable Rg between 31 and 33 Å, indicating a looser and less compact overall conformation. Backbone root mean square fluctuations (RMSF) were also evaluated to assess local flexibility (Fig. 4E,

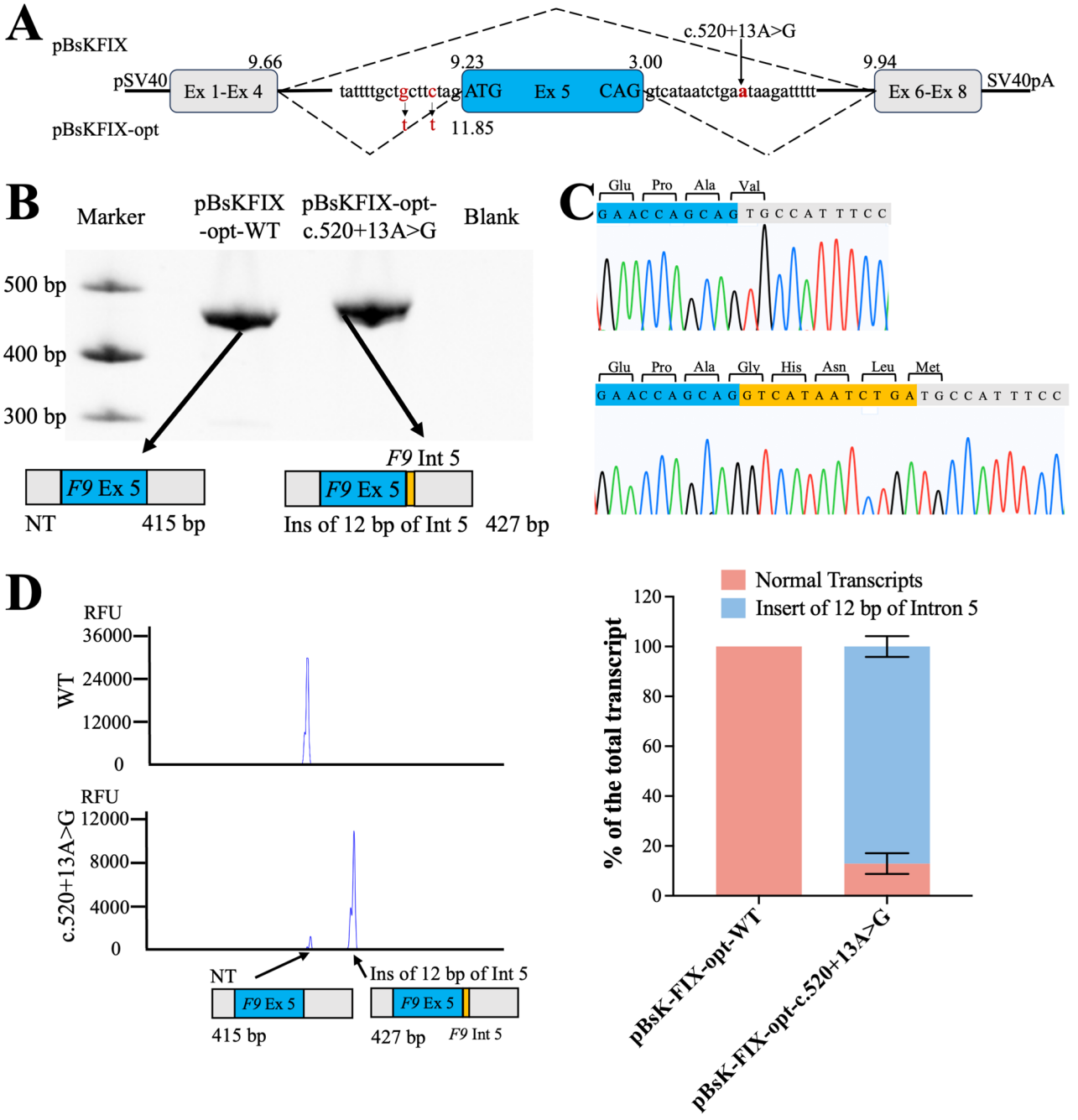


Fig. 2 Splicing analysis of the *F9* c.520 + 13 A > G variant using full-length splicing-competent constructs. **A** Schematic representation of the full-length splicing-competent cassette of the pBSK-FIX-OPT-WT vector. Exonic and intronic *F9* sequences are depicted as boxes and lines, respectively. Dotted lines indicate the major splicing pattern of exon 5 observed in the pBSK-FIX-OPT-WT construct. The scores of the 5′ and 3′ splice sites, calculated using the MaxEntScan (MES) algorithm, are reported. The pSV40 promoter (pSV40) and SV40 polyadenylation signal (SV40pA) are also shown. For additional details on the pBSK-FIX vector, refer to the Materials and Methods. **B** Gel electrophoresis of RT-PCR products obtained from HEK293T cells transiently transfected with the pBSK-FIX-opt-WT or pBSK-FIX-opt-c.520 + 13 A > G construct. **C** Sanger sequencing of RT-PCR amplicons derived from wild-type and c.520 + 13 A > G variant constructs. Twelve nucleotides from the 5′ side of intron 5 were aberrantly inserted into the transcript generated by the c.520 + 13 A > G variant. **D** Capillary electrophoresis analysis of *F9* transcripts (representative of three independent experiments). The relative proportion of each transcript was calculated based on the corresponding peak area. Product length (bp) and relative fluorescence units (RFUs) are indicated on the x- and y-axes, respectively. Data are presented as mean ± SD ($n = 3$).

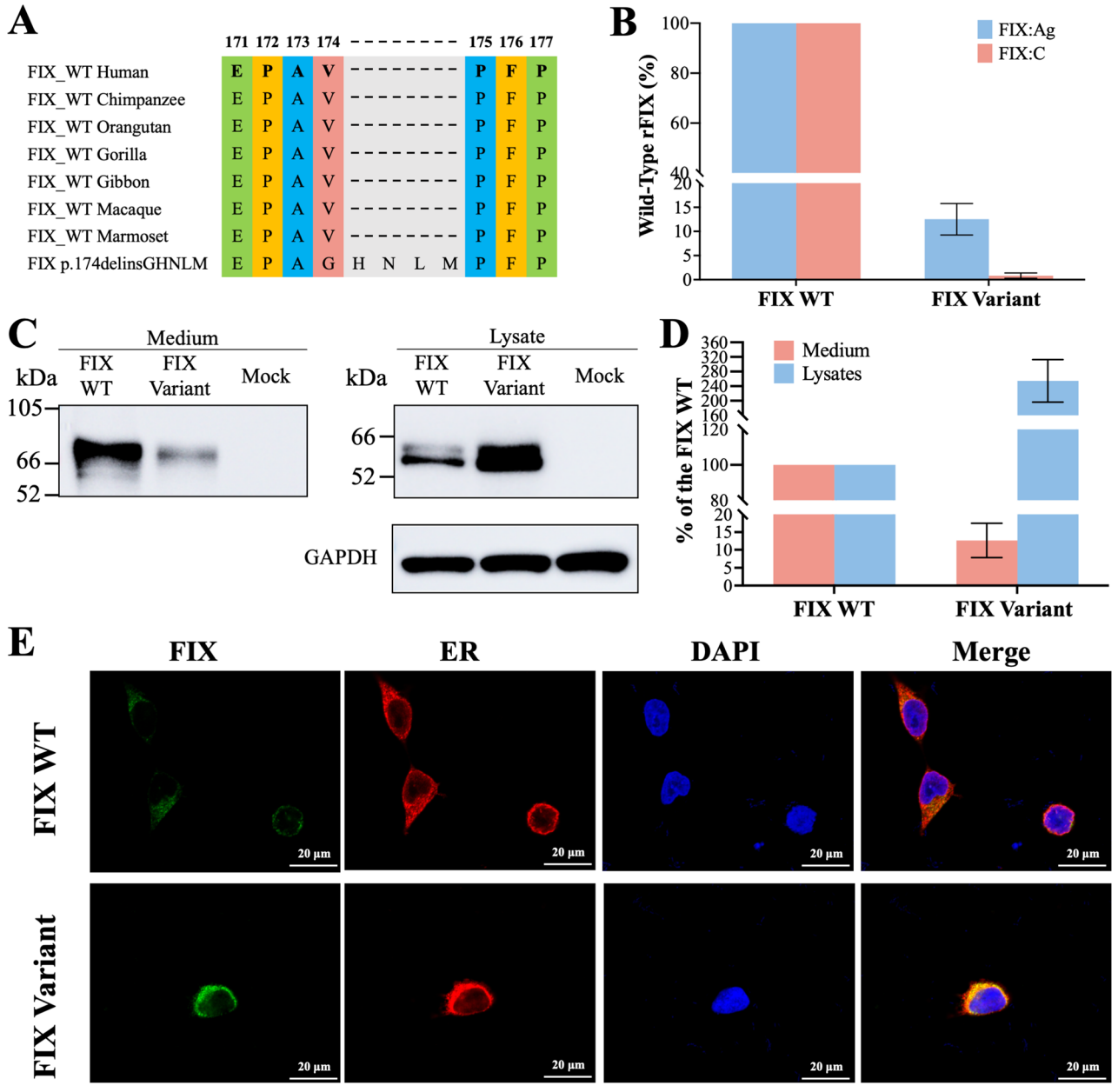


Fig. 3 The impact of rFIX p.V174delinsGHNLM variant on expression and activity of FIX. **A** Alignment of the sequences of coagulation factor IX. The amino acid sequences of 7 species were obtained from UniProt, including human, chimpanzee, orangutan, gorilla, gibbon, macaque, and marmoset. **B** Secreted FIX antigen (FIX: Ag) and activity (FIX: C) levels of rFIX p.V174delinsGHNLM variant expressed as % of rFIXwt. **C** Western blot analysis of FIX synthesis and secretion in medium and lysates of cells transfected with WT rFIX and rFIX p.V174delinsGHNLM variant. Expression plasmids were based on the pcDNA3.1(+) vector containing human FIX cDNA; the variant construct was generated from the wild-type plasmid by site-directed mutagenesis and verified by sequencing. The two bands observed in the lysate represent different glycosylation states of FIX, both of which were used for the quantification in (D). **D** The levels of rFIX p.V174delinsGHNLM secreted into the conditioned medium and expressed in the cell lysate were measured by densitometry of the bands on the Western blot and compared with those of rFIX WT. Data were obtained from three independent experiments, and error bars represent the mean ± SD. **E** The subcellular localization of FIX p.V174delinsGHNLM in transfected HEK293T cells. The HEK293T cells expressing FIX were stained with fluorescence-labeled anti-FIX antibodies (green), and the endoplasmic reticulum (ER) was marked by D3E6 antibody (red), and the nuclei were stained with DAPI (blue).

F). The RMSF profile showed that the insertion reduced fluctuations in the loop spanning residues 174–178, a region normally characterized by high flexibility, while other regions remained largely unaffected. This localized rigidity, together with the overall decrease in compactness, may perturb the conformational dynamics of FIX, potentially contributing to misfolding and increased intracellular accumulation observed in vitro. Reduced

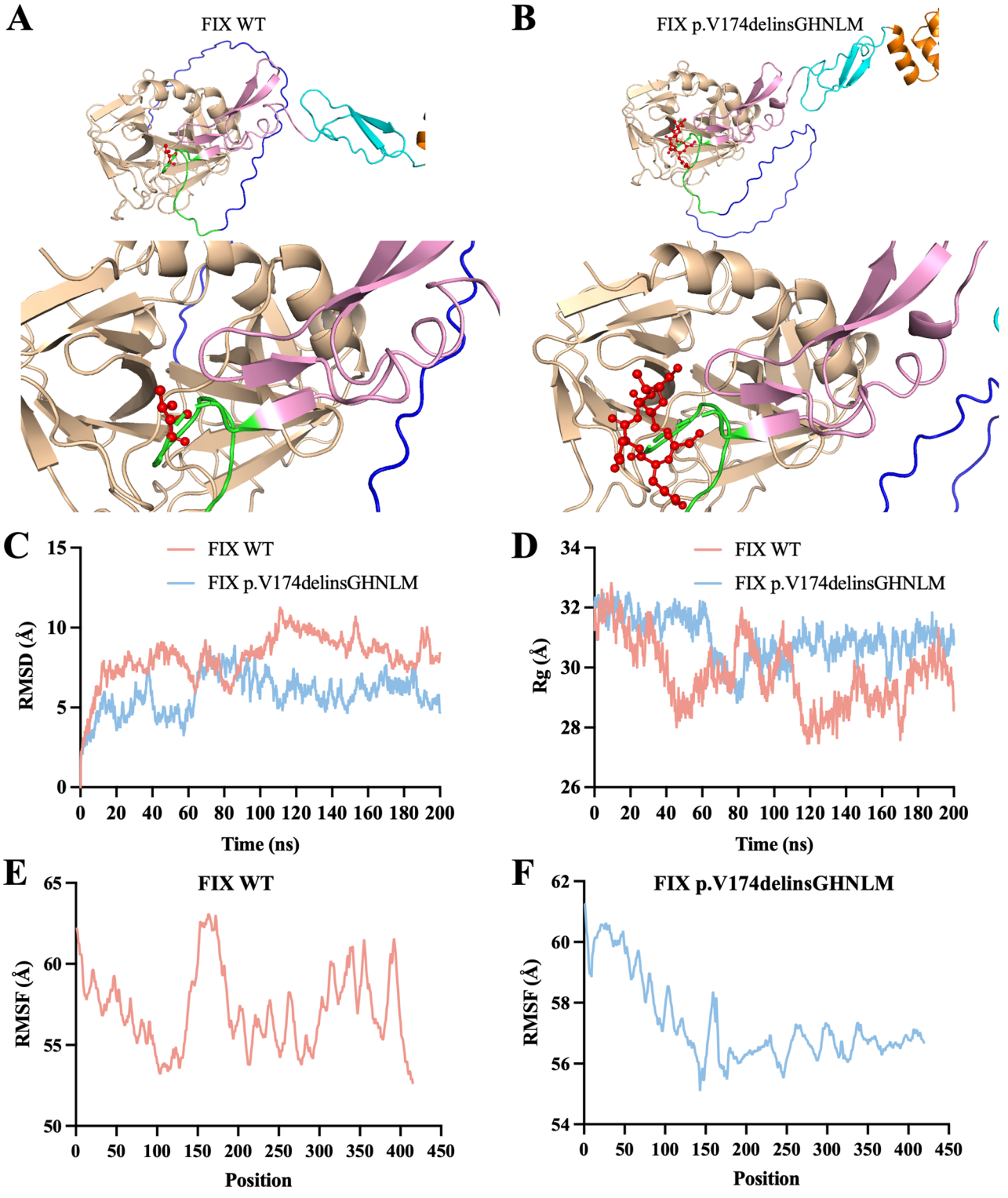


Fig. 4 Comparison of the protein structure between the FIX WT and FIX p.V174delinsGHNLM variant. Ribbon diagram of the predicted WT FIX (**A**) and p.(V174delinsGHNLM) variant (**B**) structures. The Gla domain, Epidermal Growth Factor Receptor 1 and 2 (EGF1 and EGF2) domains, activation peptide (AP), and serine protease (SP) are colored differently. The variant site p.V174 in the FIX WT model (**A**) and FIX p.V174delinsGHNLM model (**B**) are shown as ball-and-stick models. Images were generated using PyMOL. (**C**) The root mean squared deviations (**RMSD**) of the backbone atoms of the FIX models. (**D**) Radius of Gyration (**Rg**) of FIXWT and FIX insert Proteins during 200-nanosecond Molecular Dynamics Simulation. The root mean square fluctuation (RMSF) of the residue of the backbone atoms in FIX WT model (**E**) and FIX p.V174delinsGHNLM model (**F**).

flexibility in this linker region could also impair proper secretion and coagulant function.

These structural alterations likely contribute to functional abnormalities and intracellular accumulation, with decreased linker region flexibility impairing FIX coagulation activity.

Rescue of splicing by *F9* Exon5 U1snRNA results in FIX biosynthesis and coagulant activity

To investigate whether engineered snRNAs could correct the aberrant splicing caused by the *F9* c.520 + 13 A > G variant, we first tested a tailored antisense U7snRNA (pU7c.520 + 13G) designed to mask the cryptic 5′ splice site at c.520 + 12 (Fig. 5A). Co-expression of pU7c.520 + 13G with the variant construct in HEK293T cells did not increase normal transcript levels, indicating that masking the cryptic splice site alone was insufficient to redirect splicing toward the authentic 5′ splice site (Fig. 5B). These results highlight that simple blocking of the cryptic site cannot overcome the dominant effect of the variant on exon 5 inclusion.

To further explore rescue strategies, we tested an engineered U1snRNA variant (pU1Ex5) designed to specifically complement the authentic 5′ss (Fig. 5A). Co-expression with pU1Ex5 successfully restored exon inclusion, with correctly spliced transcripts representing 40% of all transcripts (Fig. 5B, C). To ensure that the rescue effect was specific to the engineered U1snRNA and not due to non-specific effects of U1 overexpression, we performed control co-transfections using pBSKFIX-opt-WT and c.520 + 13 A > G reporters with wild-type U1 snRNA. Both controls showed normal exon 5 splicing, confirming that the splicing rescue observed with pU1Ex5 is specific to the c.520 + 13 A > G variant (Figure S2). As a result, secreted FIX levels and coagulant activity increased to over 30% of normal. Western blot analysis confirmed enhanced secretion of FIX into the culture medium (Fig. 5D), while ELISA and coagulation assays further demonstrated restoration of FIX antigen levels and functional activity (Fig. 5E). These findings demonstrate that U1snRNA-mediated splicing correction efficiently restores the biosynthesis of a functional FIX.

Classification of *F9* c.520 + 13 A > G variant according to the ClinGen *F9* guideline

In line with the ClinGen *F9* guideline, the *F9* c.520 + 13 A > G variant is classified as pathogenic for HB based on several lines of evidence, including meeting the PS4_Very Strong criterion (≥ 8 probands reported) and the supporting evidence from the SpliceAI prediction tool (delta score = 0.92). However, the precise pathogenic mechanism was not fully elucidated in the ClinGen guideline. Our in vitro functional studies not only confirmed the pathogenic impact of this variant on gene function but also provided strong supporting evidence for the splicing disruption caused by this variant, which aligns with the PS3 criterion for functional studies, further strengthening its pathogenic classification.

Discussion

The *F9* c.520 + 13 A >G variant, located outside the consensus 5′ splice site, is a non-canonical splice site (NCSS) variant. Its pathogenicity is supported by both clinical and molecular observations. Documented in the dbSNP database (rs1603265507), the variant has an estimated allele frequency of G = 0.000008 (2/264690, TOPMED). The *F9* c.520 + 13 A >G variant has been reported in multiple unrelated HB patients across different countries, providing strong evidence of its pathogenicity. The study conducted by P.M. Green et al. bioinformatically predicted the impact of *F9* c.520 + 13 A >G variant on mRNA splicing [17], showing that the variant disrupts normal mRNA splicing by creating a cryptic 5′ss in intron 5, potentially leading to aberrant splicing.

However, in silico predictions often fail to accurately predict the effects of NCSS variants [32], in particular for the quantitative aspect of the aberrant splicing. Pedro Barbosa et al. [33] curated a diverse set of pathogenic deep intronic variants affecting splicing, and they found that intronic pathogenic variants located beyond 10 bp from canonical splice sites are poorly predicted. According to the ACMG criteria [34], positive in silico predictions are classified as only "supporting pathogenic" evidence, insufficient for establishing disease causality in NCSS variants. Splicing assays are therefore highly recommended to understand the pathogenic effects of NCSS variants and improve clinical interpretation. Splicing assays using splicing-competent FIX expression constructs, widely used in FIX and other coagulation factor analyses [27, 28, 35], offer a valuable alternative for assessing pre-mRNA splicing, especially given the restricted expression of *F9* in non-clinical tissues [36]. Consistent with this, we performed our analyses in HEK293T cells, which are well established for studying splicing mechanisms and recombinant FIX expression [37]. This model provides reliable and reproducible results comparable to those obtained in liver-derived cells, while avoiding interference from endogenous FIX expression. Given that exon 5 of *F9* is a weakly defined exon with a suboptimal 3′ splice site, we generated optimized constructs by selectively strengthening the 3′ss to ensure robust exon definition and allow accurate quantification of splicing outcomes in both wild-type and c.520 + 13 A >G contexts.

Here, our study dissected the molecular mechanisms underlying the frequent *F9* c.520 + 13 A >G variant, showing that the mutation severely impairs the *F9* mRNA splicing process, resulting in the formation of a defective

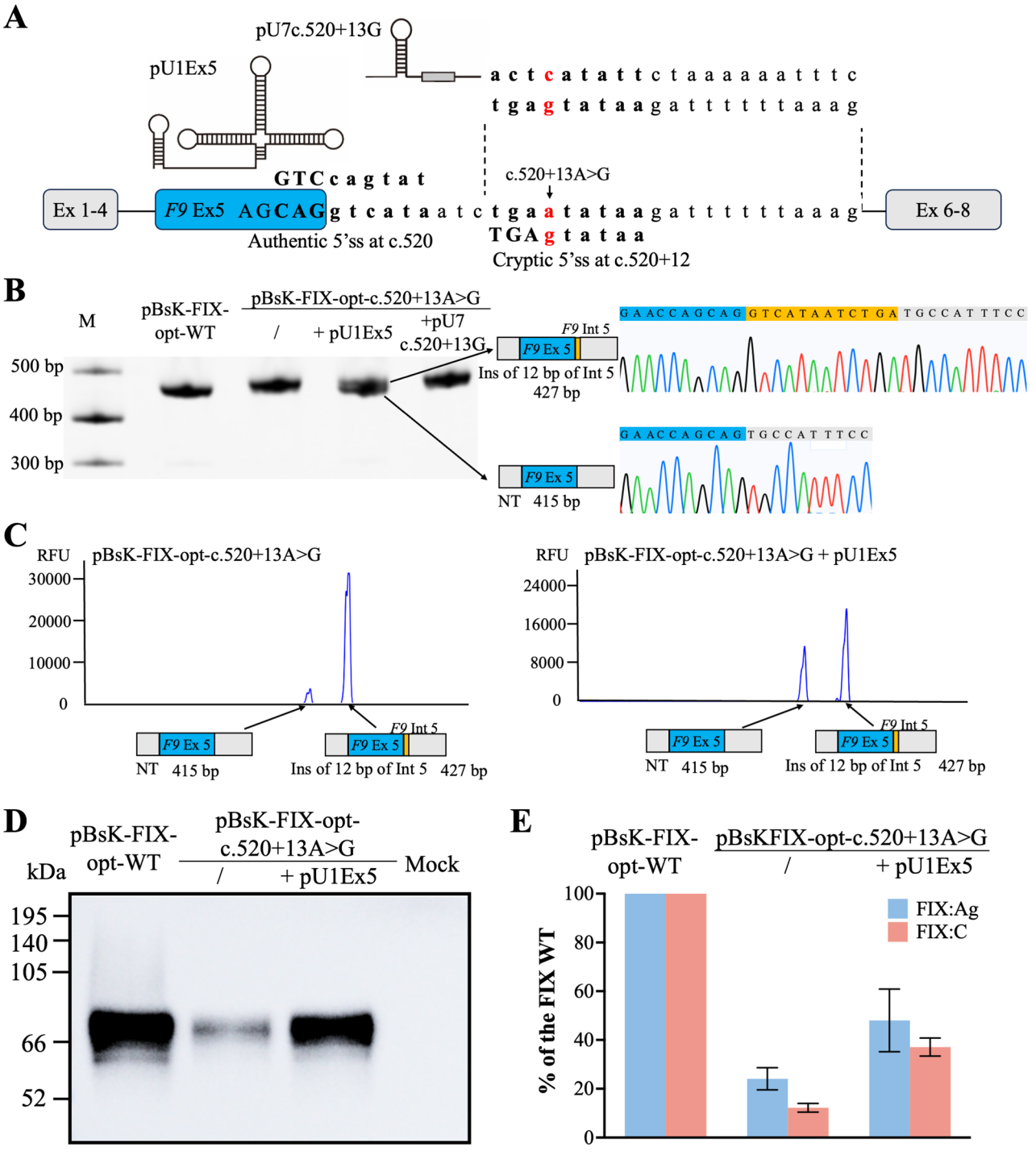


Fig. 5 Splicing rescue mediated by engineered U1snRNA and U7snRNA using full-length splicing-competent FIX constructs. **A** Schematic representation of engineered U1snRNAs (pU1Ex5) and U7snRNA (pU7c.520 + 13G) exploited in this study. **B** Analysis of *F9* exon 5 splicing patterns from HEK293T cells co-transfected with pBSK-FIX-OPT-WT or pBSK-FIX-OPT-c.520 + 13 A > G constructs alone or in combination with either pU1Ex5 or pU7c.520 + 13G. RT-PCR products were resolved on a 10% PAGE gel. M, molecular size marker (1 kb). Schematic representations of splicing isoforms are shown alongside the corresponding sequencing chromatograms of the amplicons. **C** Results of denaturing capillary electrophoresis of fluorescently labeled RT-PCR amplified fragments are shown below. The product length (bp) and relative fluorescent units (FUs) were indicated on the *x*- and *y*-axes, respectively. **D** Detection of pBSK-FIX-OPT-WT and pBSK-FIX-OPT-variant in conditioned medium by Western blot. **E** Histograms report the quantification of secreted FIX antigen (FIX: Ag) and FIX activity (FIX: C) levels of pBSK-FIX-OPT-variant alone (/) or in combination (+) with a molar excess (1.5 ×) of pU1Ex5. Results are expressed as mean ± SD of three independent experiments.

in-frame FIX (p.V174delinsGHNLM). Importantly, the molecular dissection of this frequent variant facilitated the development of tailored correction approaches based on engineered U1snRNA [38]. In particular, capillary electrophoresis enabled precise quantitative and qualitative analysis of these splicing products. The high phenotypic variability among individuals with the same variant may be attributed to complex molecular mechanisms, such as mis-splicing [39]. This is exemplified by the *F9* c.520 + 13 A >G variant in our study. Individuals with this variant exhibit significant variability in their clinical features. For example, FIX: C ranges from 1% to 20% in these HB individuals. This variability may be due to differences in the proportion of normally spliced transcripts among different individuals, which would directly affect residual FIX function.

The aberrant *F9* c.520_521insGTCATAATCTGA transcript encodes the in-frame variant FIX p.(V174delinsGHNLM). In vitro studies showed a significant decrease in secreted rFIX levels (about 10% of rFIX-wt), accompanied by an increase in intracellular rFIX levels. Further analysis revealed that the in-frame FIX (p.V174delinsGHNLM) shows impaired secretion and increased intracellular accumulation, affecting its coagulant function. Amino acid changes can have pleiotropic effects, not only affecting secretion but also frequently impairing function, which is especially critical for proteins with enzymatic activity. In fact, the combination of reduced secretion and functional impairment determines the coagulation phenotype in patients [40]. The specific coagulant activity of the secreted variant was barely detectable, making it unsuitable for therapeutic strategies based on promoting intracellular protein processing and secretion, such as with chaperone-like compounds [41, 42].

The in-frameshift variant in this study leads to amino acid insertions, disrupting both the structure and function of the protein [43]. This variant, located in the linker region of FIX, impairs normal synthesis, secretion, and coagulation activity of FIX. The linker region is essential for maintaining structural stability and proper coagulation function [44]. Molecular dynamics simulations suggest that it significantly impacts the activation peptide (AP) region. Structural changes predominantly occur in the AP region, which is crucial for FIX activation via cleavage at Arg191-Ala192 and Arg226-Val227 [45, 46]. This variant reduces the flexibility of the AP domain, potentially hindering FIX activation. Francesco Bernardi et al. [47] confirmed that variants at positions R191 and R226 influence the severity of HB, contributing to the variability in clinical HB phenotypes. This may provide an explanation from another aspect for the pleiotropic plasma phenotypes observed in HB patients carrying the *F9* c.520 + 13 A >G variant. It is important to note that AlphaFold 3 predictions, while valuable, have limitations in predicting flexible regions like the linker. Our conclusions are mainly based on RMSD and Rg, which assess overall protein stability. RMSD compares structural differences, and Rg reflects protein compactness, supporting our findings that the FIX-WT and FIX-insert variants show significant structural differences. However, further experimental validation is needed. Tools like AlphaFold3 cannot fully capture factors such as post-translational modifications. Circular dichroism (CD) or differential scanning calorimetry (DSC) could provide additional insights into the folding and stability of the WT and variant proteins.

Understanding the molecular mechanisms underlying pathological variants can be exploited to develop tailored RNA therapeutics. Engineered U1snRNA and U7snRNA have extensively demonstrated their ability to rescue splicing by promoting exon definition or masking alternative splice sites or regulatory elements [22, 48], respectively. In this study, the *F9* intronic variant c.520 + 13 A >G was efficiently rescued by a compensatory U1snRNA, confirming previous findings that an engineered U1snRNA can promote exon 5 definition [49]. In our study, the ability of a compensatory U1snRNA variant to strengthen the authentic 5′ss and thus rescue the splicing pattern further demonstrated the pathogenic role of the nucleotide change, offering a correction strategy with potential therapeutic implications for the c.520 + 13 A >G variant. Indeed, if translated into patients, the rescued transcripts would account for FIX levels well beyond the therapeutic threshold. By contrast, U7snRNA, designed to mask the cryptic splice site, did not restore correct splicing in our model. This is consistent with previous observations that U7 efficacy can be variable and context-dependent [22], with limitations arising from potential off-target binding or insufficient suppression of a strong cryptic splice site [37]. These findings underscore that, while U7snRNA remains a valuable tool in some settings [50], compensatory U1snRNA was the more effective strategy for the *F9* c.520 + 13 A >G variant. Future studies could explore combinatorial or optimized U7/U1 approaches to further enhance splicing correction efficiency and broaden therapeutic applicability.

While our study demonstrates that engineered U1snRNA can efficiently rescue the aberrant splicing caused by the *F9* c.520 + 13 A >G variant in cell-based assays, translating this approach into a therapeutic setting will require addressing delivery and durability. Efficient and tissue-specific delivery of U1snRNA to hepatocytes, the natural site of FIX synthesis, remains a critical challenge. Advances in RNA therapeutics, such as adeno-associated viral (AAV) vectors [51], lipid nanoparticles (LNPs) [52], and chemically modified oligonucleotides, provide promising platforms to achieve sustained expression

and targeted delivery of engineered snRNAs. Indeed, Dario Balestra and colleagues have demonstrated in vivo rescue of a splicing-defective human *F7* variant (c.859 + 5G >A) in a murine model using plasmid and AAV-mediated delivery of engineered U1snRNA, leading to restored plasma FVII levels and correct hepatic splicing [35]. These results provide a proof-of-principle that U1snRNA can be effectively delivered in vivo. However, further evaluation is needed to establish long-term safety, immune compatibility, and therapeutic efficacy in preclinical models of hemophilia B. By combining molecular precision with clinically validated RNA delivery systems, engineered U1snRNAs may represent a realistic and personalized therapeutic strategy for patients carrying the *F9* c.520 + 13 A >G variant.

In addition to RNA-based approaches, genome editing represents an alternative or complementary strategy for correcting the *F9* c.520 + 13 A >G variant. Targeted disruption of the cryptic 5′ splice site using CRISPR-Cas systems, base editors (BE) [53], or prime editors (PE) [54] could restore correct splicing and normal FIX production in hepatocytes. While this approach offers the advantage of potentially permanent correction, it also faces challenges, including efficient delivery to liver cells and minimization of off-target effects. Future studies applying these genome editing strategies in relevant models could provide valuable insights into long-term correction of splicing variants in hemophilia B.

In conclusion, through bioinformatic, splicing and expression assays, we demonstrated that the *F9* c.520 + 13 A > G variant is mainly associated with aberrant mRNA splicing, leading to the production of the variant FIX p.(V174delinsGHNLM), which exhibits extracellular secretory defects and no coagulation activity. Overall, this study elucidates the molecular mechanisms underlying the *F9* c.520 + 13 A > G variant and provides evidence of the correction potential of a tailored compensatory U1snRNA, thus representing a personalized therapeutic strategy for HB.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40246-025-00890-y.

Supplementary Material

Acknowledgements
We thank Dr. Emanuele Buratti and Dr. Franco Pagani for generously providing the U1snRNA, and U7snRNA plasmids.

Author contributions
Huayang Zhang performed experiments, analyzed data, and drafted the manuscript. Chong Wang and Meixiu Gu contributed to experimental procedures and data interpretation. Zhimin Meng, Yichao Guo, and Weitao Zhang provided technical support and assisted with data analysis. Dario Balestra designed the study, provided critical feedback, and contributed to manuscript revisions. Wei Guo and Beili Wang supervised the project, and approved the final manuscript. All authors reviewed the manuscript and agreed with its content.Huayang Zhang and Chong Wang contributed equally as the first authors.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Laboratory Medicine, Zhongshan Hospital, Fudan University, 180 Fenglin Road, Shanghai, China
[2]Department of Life Science and Biotechnology, University of Ferrara, Ferrara, Italy
[3]Department of Laboratory Medicine, Shanghai Geriatric Medical Center, Shanghai, China
[4]Department of Laboratory Medicine, Wusong Branch, Zhongshan Hospital, Fudan University, Shanghai, China
[5]Department of Laboratory Medicine, Xiamen Branch, Zhongshan Hospital, Fudan University, Xiamen,, China

Received: 19 April 2025 / Accepted: 3 December 2025
Published online: 08 January 2026

References

1. Peyvandi F, Garagiola I, Young G. The past and future of haemophilia: diagnosis, treatments, and its complications. Lancet. 2016;388(10040):187–97. https://doi.org/10.1016/s0140-6736(15)01123-x.
2. Bolton-Maggs PH, Pasi KJ, Haemophilias A. B Lancet. 2003;361(9371):1801–9. https://doi.org/10.1016/s0140-6736(03)13405-8.
3. Lillicrap D. The molecular basis of haemophilia B. Haemophilia. 1998;4(4):350–7. https://doi.org/10.1046/j.1365-2516.1998.440350.x.
4. Hsu TC, Nakaya SM, Thompson AR. Severe haemophilia B due to a 6 kb factor IX gene deletion including exon 4: non-homologous recombination associated with a shortened transcript from whole blood. Thromb Haemost. 2007;97(2):176–80.
5. Balestra D, Barbon E, Scalet D, et al. Regulation of a strong F9 cryptic 5′ss by intrinsic elements and by combination of tailored U1snRNAs with antisense oligonucleotides. Hum Mol Genet. 2015;24(17):4809–16. https://doi.org/10.1093/hmg/ddv205.
6. Heit JA, Ketterling RP, Zapata RE, Ordonez SM, Kasper CK, Sommer SS. Haemophilia B Brandenberg-type promoter mutation. Haemophilia. 1999;5(1):73–5. https://doi.org/10.1046/j.1365-2516.1999.00193.x.
7. Kaufman RJ. Post-translational modifications required for coagulation factor secretion and function. Thromb Haemost. 1998;79(6):1068–79.
8. Simhadri VL, Hamasaki-Katagiri N, Lin BC, et al. Single synonymous mutation in factor IX alters protein properties and underlies haemophilia B. J Med Genet. 2017;54(5):338–45. https://doi.org/10.1136/jmedgenet-2016-104072.
9. Wu PC, Hamaguchi N, Yu YS, Shen MC, Lin SW. Hemophilia B with mutations at glycine-48 of factor IX exhibited delayed activation by the factor VIIa-tissue factor complex. Thromb Haemost. 2000;84(4):626–34.
10. Xu Z, Spencer HJ, Harris VA, Perkins SJ. An updated interactive database for 1692 genetic variants in coagulation factor IX provides detailed insights into hemophilia B. J Thromb Haemost. 2023;21(5):1164–76. https://doi.org/10.1016/j.jtha.2023.02.005.
11. Vaz-Drago R, Custódio N, Carmo-Fonseca M. Deep intronic mutations and human disease. Hum Genet. 2017;136(9):1093–111. https://doi.org/10.1007/s00439-017-1809-4.
12. Dericquebourg A, Fretigny M, Chatron N, et al. Whole F9 gene sequencing identified deep intronic variations in genetically unresolved hemophilia B patients. J Thromb Haemost. 2023;21(4):828–37. https://doi.org/10.1016/j.jtha.2022.12.005.
13. Belvini D, Salviato R, Radossi P, et al. Molecular genotyping of the Italian cohort of patients with hemophilia B. Haematologica. 2005;90(5):635–42.
14. Koeberl DD, Bottema CD, Ketterling RP, Bridge PJ, Lillicrap DP, Sommer SS. Mutations causing hemophilia B: direct estimate of the underlying rates of

spontaneous germ-line transitions, transversions, and deletions in a human gene. Am J Hum Genet. 1990;47(2):202–17.
15. Li X, Drost JB, Roberts S, Kasper C, Sommer SS. Factor IX mutations in South Africans and African Americans are compatible with primarily endogenous influences upon recent germline mutations. Hum Mutat. 2000;16(4):371. https://doi.org/10.1002/1098-1004(200010)16:4%3C371::Aid-humu11%3E3.0.Co;2-p.
16. Parrado Jara YA, Yunis Hazbun LK, Linares A, Yunis Londoño JJ. Molecular characterization of hemophilia B patients in Colombia. Mol Genet Genomic Med. 2020;8(5):e1210. https://doi.org/10.1002/mgg3.1210.
17. Anagnostopoulos T, Morris AP, Ayres KL, Giannelli F, Green PM. DNA variation in a 13-Mb region including the F9 gene: inferring the genealogical history and causal role of a hemophilia B mutation (IVS 5 + 13 A–>G). J Thromb Haemost. 2003;1(12):2609–14. https://doi.org/10.1111/j.1538-7836.2003.00514.x.
18. Pinotti M, Balestra D, Rizzotto L, Maestri I, Pagani F, Bernardi F. Rescue of coagulation factor VII function by the U1 + 5A SnRNA. Blood. 2009;113(25):6461–4. https://doi.org/10.1182/blood-2009-03-207613.
19. Buratti E, Chivers M, Královicová J, et al. Aberrant 5′ splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. Nucleic Acids Res. 2007;35(13):4250–63. https://doi.org/10.1093/nar/gkm402.
20. Schümperli D, Pillai RS. The special Sm core structure of the U7 snrnp: far-reaching significance of a small nuclear ribonucleoprotein. Cell Mol Life Sci. 2004;61(19–20):2560–70. https://doi.org/10.1007/s00018-004-4190-0.
21. Gadgil A, Raczyńska KD. U7 snrna: A tool for gene therapy. J Gene Med. 2021;23(4):e3321. https://doi.org/10.1002/jgm.3321.
22. Ferraresi P, Balestra D, Guittard C, et al. Next-generation sequencing and Recombinant expression characterized aberrant splicing mechanisms and provided correction strategies in factor VII deficiency. Haematologica. 2020;105(3):829–37. https://doi.org/10.3324/haematol.2019.217539.
23. Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37(9):e67. https://doi.org/10.1093/nar/gkp215.
24. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biology J Comput Mol Cell Biology. 2004;11(2–3):377–94. https://doi.org/10.1089/1066527041410418.
25. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in genie. Journal Comput Biology J Comput Mol Cell Biology Fall. 1997;4(3):311–23. https://doi.org/10.1089/cmb.1997.4.311.
26. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting splicing from primary sequence with deep learning. Cell. 2019;176(3):535–e54824. https://doi.org/10.1016/j.cell.2018.12.015.
27. Balestra D, Scalet D, Pagani F, et al. An Exon-Specific U1snRNA induces a robust factor IX activity in mice expressing multiple human FIX splicing mutants. Mol Ther Nucleic Acids. 2016;5(10):e370. https://doi.org/10.1038/mtna.2016.77.
28. Scalet D, Maestri I, Branchini A, Bernardi F, Pinotti M, Balestra D. Disease-causing variants of the conserved + 2T of 5′ splice sites can be rescued by engineered U1snRNAs. Hum Mutat. 2019;40(1):48–52. https://doi.org/10.1002/humu.23680.
29. Zhang H, Chen C, Wu X, et al. Effects of 14 F9 synonymous codon variants on hemophilia B expression: alteration of splicing along with protein expression. Hum Mutat. 2022;43(7):928–39. https://doi.org/10.1002/humu.24377.
30. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. Nature. 2024;630(8016):493–500. https://doi.org/10.1038/s41586-024-07487-w.
31. 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(7):1329–40. https://doi.org/10.1111/jth.12276.
32. Lord J, Baralle D. Splicing in the diagnosis of rare disease: advances and challenges. Front Genet. 2021;12:689892. https://doi.org/10.3389/fgene.2021.689892.
33. Barbosa P, Savisaar R, Carmo-Fonseca M, Fonseca A. Computational prediction of human deep intronic variation. Gigascience. 2022;12. https://doi.org/10.1093/gigascience/giad085.
34. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. Genet Med. 2015;17(5):405–24. https://doi.org/10.1038/gim.2015.30.
35. Balestra D, Faella A, Margaritis P, et al. An engineered U1 small nuclear RNA rescues splicing defective coagulation F7 gene expression in mice. J Thromb Haemost. 2014;12(2):177–85. https://doi.org/10.1111/jth.12471.
36. Green PM, Rowley G, Giannelli F. Unusual expression of the F9 gene in peripheral lymphocytes hinders investigation of F9 mRNA in hemophilia B patients. J Thromb Haemost. 2003;1(12):2675–6. https://doi.org/10.1111/j.1538-7836.2003.0543a.x.
37. Zhang H, Xin M, Lin L, Chen C, Balestra D, Ding Q. Pleiotropic effects of different exonic nucleotide changes at the same position contribute to hemophilia B phenotypic variation. J Thromb Haemost. 2024;22(4):975–89. https://doi.org/10.1016/j.jtha.2023.12.031.
38. Baralle D, Buratti E. RNA splicing in human disease and in the clinic. Clin Sci (Lond). 2017;131(5):355–68. https://doi.org/10.1042/cs20160211.
39. Hull J, Campino S, Rowlands K, et al. Identification of common genetic variation that modulates alternative splicing. PLoS Genet. 2007;3(6):e99. https://doi.org/10.1371/journal.pgen.0030099.
40. Branchini A, Campioni M, Mazzucconi MG, et al. Replacement of the Y450 (c234) phenyl ring in the carboxyl-terminal region of coagulation factor IX causes pleiotropic effects on secretion and enzyme activity. FEBS Lett. 2013;587(19):3249–53. https://doi.org/10.1016/j.febslet.2013.08.019.
41. Andersen E, Chollet ME, Baroni M, et al. The effect of the chemical chaperone 4-phenylbutyrate on secretion and activity of the p.Q160R missense variant of coagulation factor FVII. Cell Biosci. 2019;9:69. https://doi.org/10.1186/s13578-019-0333-8.
42. Pignani S, Todaro A, Ferrarese M, et al. The chaperone-like sodium phenylbutyrate improves factor IX intracellular trafficking and activity impaired by the frequent p.R294Q mutation. J Thromb Haemost. 2018;16(10):2035–43. https://doi.org/10.1111/jth.14236.
43. Montgomery SB, Goode DL, Kvikstad E, et al. The origin, evolution, and functional impact of short insertion-deletion variants identified in 179 human genomes. Genome Res. 2013;23(5):749–61. https://doi.org/10.1101/gr.148718.112.
44. 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 USA. 1995;92(21):9796–800. https://doi.org/10.1073/pnas.92.21.9796.
45. Zögg T, Brandstetter H. Activation mechanisms of coagulation factor IX. Biol Chem. 2009;390(5–6):391–400. https://doi.org/10.1515/bc.2009.057.
46. Goodeve AC. Hemophilia B: molecular pathogenesis and mutation analysis. J Thromb Haemost. 2015;13(7):1184–95. https://doi.org/10.1111/jth.12958.
47. Branchini A, Morfini M, Lunghi B, et al. F9 missense mutations impairing factor IX activation are associated with pleiotropic plasma phenotypes. J Thromb Haemost. 2022;20(1):69–81. https://doi.org/10.1111/jth.15552.
48. Romano G, Riccardi F, Bussani E, et al. Rescue of a Familial Dysautonomia mouse model by AAV9-Exon-specific U1 SnRNA. Am J Hum Genet. 2022;109(8):1534–48. https://doi.org/10.1016/j.ajhg.2022.07.004.
49. Tajnik M, Rogalska ME, Bussani E, et al. Molecular basis and therapeutic strategies to rescue factor IX variants that affect splicing and protein function. PLoS Genet. 2016;12(5):e1006082. https://doi.org/10.1371/journal.pgen.1006082.
50. Krooss S, Werwitzke S, Kopp J, et al. Pathological mechanism and antisense oligonucleotide-mediated rescue of a non-coding variant suppressing factor 9 RNA biogenesis leading to hemophilia B. PLoS Genet. 2020;16(4):e1008690. https://doi.org/10.1371/journal.pgen.1008690.
51. Donadon I, Bussani E, Riccardi F, et al. Rescue of spinal muscular atrophy mouse models with AAV9-Exon-specific U1 SnRNA. Nucleic Acids Res. 2019;47(14):7618–32. https://doi.org/10.1093/nar/gkz469.
52. Garaizar A, Díaz-Oviedo D, Zablowsky N, et al. Toward Understanding lipid reorganization in RNA lipid nanoparticles in acidic environments. Proc Natl Acad Sci USA. 2024;121(45):e2404555121. https://doi.org/10.1073/pnas.2404555121.
53. Peretto L, Tonetto E, Maestri I, et al. Counteracting the common Shwachman-Diamond Syndrome-Causing SBDS c.258 + 2T > C mutation by RNA therapeutics and Base/Prime editing. Int J Mol Sci. 2023;24(4). https://doi.org/10.3390/ijms24044024.

54. Tonetto E, Cucci A, Follenzi A, Bernardi F, Pinotti M, Balestra D. DNA base editing corrects common hemophilia A mutations and restores factor VIII expression in in vitro and ex vivo models. J Thromb Haemost. 2024;22(8):2171–83. https://doi.org/10.1016/j.jtha.2024.04.020.

Publisher's note