## Supplemental information for

# A Systematic Approach for Evaluating the Role of Surface-Exposed Loops in Trypsin-like Serine Proteases Applied to the 170 loop in Coagulation Factor VIIa

Anders B. Sorensen,<sup>1,2</sup> Per Jr. Greisen,<sup>1</sup> Jesper J. Madsen,<sup>3,4</sup> Jacob Lund,<sup>1</sup> Gorm Andersen,<sup>1</sup> Pernille G. Wulff-Larsen,<sup>1</sup> Anette A. Pedersen,<sup>1</sup> Prafull S. Gandhi,<sup>1</sup> Michael T. Overgaard,<sup>2</sup> Henrik Østergaard,<sup>1</sup> Ole H. Olsen<sup>1,5,\*</sup>

<sup>1</sup> Global Research, Novo Nordisk A/S, 2760 Måløv, Denmark

<sup>2</sup> Department of Chemistry and Bioscience, Aalborg University, 9220 Aalborg, Denmark

<sup>3</sup> Global and Planetary Health, College of Public Health, University of South Florida, Tampa, FL 33612, USA

<sup>4</sup> Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA

<sup>5</sup> Current affiliation: Novo Nordisk Foundation Center for Basic Metabolic Research, Section for Metabolic Receptology, University of Copenhagen, Blegdamsvej 3b, DK-2200 Copenhagen, Denmark

### **Supplemental methods**

#### FX Activation Kinetics

500 nM of individual FVIIa-variants were pre-incubated with 20 µM PS:PC (25:75 Haematologic technologies, US) in 50 mM HEPES pH 7.4, 100 mM NaCl, 10 mM CaCl2, 0.1 % PEG8000 for 5 minutes at room temperature. 500 nM FVIIA-WT in the absence and presence of 100 nM soluble tissue factor (sTF) where used for comparison. Plasma-derived FX (Haematologic technologies, US) was then added to a final concentration of 2.5 µM. At indicated time points, a small portion of the sample was withdrawn, mix with NuPAGE<sup>TM</sup> LDS Sample Buffer and NuPAGE<sup>TM</sup> Sample Reducing Agent (Thermo Fisher Scientific, US) and heated at 70 °C for 10 minutes. Samples were separated in 4-12% Bis-Tris SDS-PAGE (Thermo Fisher Scientific, US) and stained with coomassie brilliant blue. Plasma-derived FXa (Haematologic technologies, US) were added as a control. Gel band intensities (area under the curve) were evaluated by an in-house dedicated python script, using the FXa HCβ band to determine initial FX-activation rates using a first order equation in GraphPad Prism.

### This supplementary document contains the following figures:

Figure S1 Sequence alignment of the catalytic domain of selected trypsin-like proteases
Figure S2 Heatmap of functional FVIIa variant characterization with 3 μM sTF
Figure S3: RMSF plots of all reconstructed loops from FVIIa-WT (1dan)
Figure S4: REU vs RMSD scatterplots for FVIIa-WT (1dan)
Figure S5: RMSF plots of all reconstructed loops from FVIIa-Y<sub>T</sub> (4z6a)
Figure S6: REU vs RMSD scatterplots for FVIIa-Y<sub>T</sub> (4z6a)
Figure S7: Comparison of trypsin and FVIIa var. 36
Figure S8: FX-activation kinetics of selected FVIIa-variants visualized by SDS-PAGE

1					
g) loop	YVP	YNAA I YNWR - YDR (312) loop	KVGDS KAEAS HSSNT	но 9	€€∃:::::::
99c (23	// I PST // I PST // MPDK / I RHEE AFTHPA /RLHEA	Y I HPR I RHPQ 170c	LQQSR LQQSR LEHAK EKQAN VQVSN	SAPDV QKRYR LRSTK KDSTR EASYP	LLRAP LLRAP - FRLP 
	RVAQV RVTQV GVSKI QVDQN AVRSV GV2EI	MLEKI	MTQDO MTQDO MTQDO RTQEO	SLERC TNEEC DRATC ERPVC SQAKC	Р К Р 6 V V V V V V V V V V V V V V V V V V
70c (210) loop	GDEQSR GDEQSR GDEQVR KTEQES GTEQL EPCQTL EDTSFF	HTEOKR IN I EK I S GNEOF I	LNVPRL LNVPRL IEVPRL VHLPRV LLVPRI	AQVPFL AKIPLV LRVPLV VNLPLV LDAPVL	KLMRSE KLMRSE RHMDSK RHMDSK RHMDSK EKTMDSK EKTVS- EKTVS- EKTQAV KVIDQF
	E HD HSC HSC HSC HSC HSC HSC HSC HSC HSC HSC		ELMA FLMRA LLRR		
	I AVLGEHDLS T AVLGEHDLS T VVMGEHDFS RVRLGEYSVH R I VAGEHDLE T VVLGGERRN R VYSGILNOS	TVVAGEHNIE LVRIGKHSR1 QVRLGEHNIE 15c (288) loop		E E Y AS K I QN R SAL T AN VGKGQP A DYPE	P SUYTRUSQY SUYTRUSQY SUYTRUSQY SUYTRUSQY SUYTRUSSY SUYTRUSSY SUYTRUSSY SUYTRUSSY SUYTRUSSY SUYTRUSSY SUYTRUSSY SUYTRUSSY
H57c (193) 60c (190) loop	RNL GNI KIL	TENDL	LDRG LDRG LDRG LDRG SEDG	RKLRD FHKG KETW FASSG	
	NW	WDKNF	SGWGQI SGWGQI SGWGQI SGWGRI SGWGKI	AGWGHC TGWGYF SGWGRN TGWGNI SGWGNI	220 GGGCAA GGGCCAA GGGCCACAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCACAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAA GGGCCAAAA GGGCCAAAA GGGCCAAAA GGGCCAAAA GGGCCAAAAA GGGCCAAAA GGGCCAAAAA GGGCCAAAAA GGGCCAAAAAAAA
	FDKIK FDKIK FDNIR LDYAH LEKLK LQDRP FYGVE	VETG- LLYPP YK	RFSLV RFSLV RFSRV KFSMV SKHTV	TLCQV TDCWV GSGYV YKGRV TKCLI	G I V SW G I SW G I SW G I SW G I SW
	SAAHC SAAHC TAAHC TAAHC TAAHC TAAHC	TAAHO	TLAFV TLAFV TLARI ELSSI ELWAV	RPSET RNVIY IFLKF LLQAG - PATG	TWYLT TWYLT TWYLT TWFLT TMFLT TAFLL RUYLV VWHLV VWHLV VWHLV CWHLV CWHLV
35c (172) loop	IT IWVV ARWIV BPEWVV CPTWIL	DRWVL	UTFSER UTFSER SFSEN CRFAVY CRFAVY	64 А (G D (EYT - N (EYT - N (P	
	GTLIN GTLIN GYLIN GSLLS GSLLS GSLLS	GS I VN ASL I S GSL I N	CLPER CLPER CLPER CLPER	CLPSC CLPSK CLPSK CLPDK SLPTA	HATHY HATHY HATKY HATKY LVTRY LVCRD LVCCD VVCN-
	AQLCG LLLCG KGKCG KGKCG HSFCA	DAFCG DELLCG YHFCG	HVVPL HVVPL FVVPL	YVQPV SQRPI YVTPI YIHPV RVSTI	1450 (344 150 (344 150 (347 150 (347)
	NG	KV	-VLTC -VLTC -TFTC -NLTC -VYSV	ALLSF - NYTE - VLNS - AFSE - VINA	
	VLLLV AVLKLV AVLKI ALLIQ ALLKY AALYW AALYW	VVL NG		V	(334)100 
	ECPW0 ECPW0 ECPW0 AHPYI	MS PWG SVPYG	RLHQP RLQQP RLHRP KLETP RLRTP	KL GED FLDEP KL KKP KL SSR	1860 9780 - 1780 - 1780 - 1780 - 1860
inal tail	VCPKG VCPKG VCPKG TCPPG FCPKG VALRG	DAKPG DAEIG NCEEN	DIALL DIALL	DLALL DIALL DIALL DIALM	
N-tern			GTTNH GATNH GKINH GKINH GQVNH	VSYQF AESGY NKYNF ENLDF KTLNN	PNITE PNITE PNITE PNITE PNITE PNITE PNITE PNITE PNITE PNITE
	sapiens) icicularis) nusculus) gallus) .' rerio) NP_000119.1) IP_000119.1)	4P_000124.1) 1P_000497.1) _002760)	sapiens) scicularis) nusculus) gallus) '. rerio)	NP_000496.2 VP_000119.1) VP_000124.1) IP_000497.1) _002760)	sapiens) sapiens) unsculus) galus) . rerio) P_000139.1) P_000124.1) P_000124.1) P_000437.1)
	Iman ( <i>H. fas</i> no ( <i>M. fas</i> ouse ( <i>M. r</i> nicken ( <i>G.</i> brafish ( <i>D</i> ctor XIIa ( <i>i</i> ctor XIIa ( <i>i</i> )	ctor IXa (h rombin (N ypsin (NP_	Iman (H. fas no (M. fas ouse (M. n nicken (G. brafish (D	ictor XIIa ( ictor XIa ( ictor IXa († irombin (N ypsin (NP_	Iman (H. 1 Ino (M. fas uuse (M. m incken (G. brafish (D ctor XIa ( ctor XIa ( ctor XIa ( rombin (N rombin (N rombin (N
	TO ZOZOZ Lactor VIIa Eactor VIIa Eactor VIIa	ᄣᅣᅣ	Factor VIIa ± 0. ≦ Ω ⊈ I		王인호호호호 문문년 국민교교 철유호로 118

**Supplemental Figure S1.** Sequence alignment of the catalytic domain of selected trypsin-like proteases using the PROMALS3D webserver. Consensus numbering based on Chymotrypsin has been used to mark surface-exposed loops with extended loop regions denoted with capital letters. FVII numbering is shown in parentheses for reference. The red colour gradient indicates conservation level with intense colour showing the highest level of conservation.

3 uM Tissue Factor												
-	Deletion Overview	S2288 K <sub>M</sub> S2288 k <sub>cat</sub> S2		S2288 xWT	2288 xWT pABA x WT		Barrahalan C					
NO.		[mM]	[s <sup>-1</sup> ]	[%]	[%]	[%]	Remaining Seq.					
FVIIa-WT	LQQSRKVGDSPN	1.1	36.1	100.0	100.0	100.0	LQQSRKVGDSPN	14				
1	LQQSRKVGDSPN	3.7	20.3	17.3	22.5	41.1	LQSRKVGDSPN	1				
2	LQQSRKVGDSPN	5.7	19.7	10.9	13.8	30.7	LQQRKVGDSPN					
3	LQQS <mark>R</mark> KVGDSPN	1.7	29.5	55.0	100.8	0.0	LQQSKVGDSPN					
4	LQQSRKVGDSPN	1.8	27.2	48.1	101.7	67.8	LQQSRVGDSPN					
<b>→</b> 5	LQQSRKVGDSPN	1.7	26.5	50.4	94.4	86.3	LQQSRKGDSPN	H				
6	LQQSRKVGDSPN	1.3	26.3	63.3	127.8	72.9	LQQSRKVDSPN	ĩ				
7	LQQSRKVGDSPN	2.3	48.2	65.3	61.9	118.8	LQQSRKVGSPN					
8	LQQSRKVGDSPN	1.6	32.9	64.0	69.3	76.0	LQQSRKVGDPN					
9	LQQSRKVGDSPN	1.7	39.1	74.1	104.0	88.8	LQQSRKVGDSN	1				
10	LQQSRKVGDSPN	3.1	17.8	18.0	27.2	45.6	LQRKVGDSPN	Ĩ				
11	LQQSRKVGDSPN	2.7	25.7	30.1	35.0	52.1	LQQKVGDSPN					
12	LQQS <mark>RK</mark> VGDSPN	1.6	34.0	67.5	117.9	74.9	LQQSVGDSPN					
13	LQQSRKVGDSPN	2.0	33.4	52.2	88.0	84.0	LQQSRGDSPN	÷				
14	LQQSRKVGDSPN	1.7	27.2	50.5	70.2	98.4	LQQSRKDSPN	N				
15	LQQSRKVGDSPN	2.8	34.8	39.5	55.1	64.6	LQQSRKVSPN					
16	LQQSRKVGDSPN	1.6	32.3	63.0	60.1	123.5	LQQSRKVGPN					
17	LQQSRKVGDSPN	1.6	33.6	66.9	136.0	55.4	LQQSRKVGDN	1				
18	LOOSRKVGDSPN	1.8	24.6	43.8	74.0	72.0	LOKVGDSPN	Ŧ				
19	LQQSRKVGDSPN	1.5	26.3	56.0	109.1	61.1	LQQVGDSPN					
20	LQQSRKVGDSPN	5.6	27.7	15.6	27.9	36.0	LQQSGDSPN	- L				
ພ 21	LQQSRKVGDSPN	3.3	35.9	34.6	55.2	53.9	LQQSRDSPN	1				
22	LQQSRKVGDSPN	4.6	45.3	31.2	70.3	78.7	LQQSRKSPN	- Ì				
23	LQQSRKVGDSPN	4.2	27.9	21.1	59.6	66.9	LQQSRKVPN					
24	LQQSRKVGDSPN	4.4	14.9	10.9	46.1	21.5	LQQSRKVGN	1				
25	LOOSRKVGDSPN	2.6	31.9	38.9	60.5	24.6	LQVGDSPN	ĩ				
26	LQQSRKVGDSPN	2.6	8.9	10.7	68.5	10.2	LQQGDSPN					
27	LQQSRKVGDSPN	2.7	24.7	29.1	83.9	35.3	LQQSDSPN	÷				
28	LQQSRKVGDSPN	2.2	26.3	38.6	94.5	46.5	LQQSRSPN	C				
29	LQQSRKVGDSPN	5.0	4.5	2.8	26.2	3.4	LQQSRKPN					
30	LQQSRKVGDSPN	5.3	31.9	19.2	45.8	41.7	LQQSRKVN	1				
31	LOOSRKVGDSPN	1.4	8.9	19.5	90.3	14.8	LOGDSPN	ī				
32	LOOSRKVGDSPN	1.9	27.2	44.5	85.7	21.1	LOODSPN					
л 33	LOOSRKVGDSPN	2.1	6.6	9.8	87.1	8.9	LOOSSPN	9				
34	LQQSRKVGDSPN	2.3	16.3	23.0	43.4	49.0	LQQSRPN	- T				
35	LQQSRKVGDSPN	8.7	19.9	7.3	21.3	13.1	LQQSRKN	1				
36	LOOSRKVGDSPN	0.2	15.6	210.0	666.8	41.0	LODSPN	Ŧ				
א 37	LOOSRKVGDSPN	2.5	24.9	32.2	83.9	15.5	LOOSPN	0				
38	LQQSRKVGDSPN	2.0	15.2	24.3	52.4	18.0	LQQSRN	Î				
v 39	LQQSRKVGDSPN	2.5	42.6	54.2	82.8	18.3	LQQPN	1				

**Supplemental Figure S2.** Heatmap of Functional FVIIa Variant Characterization With 3  $\mu$ M Soluble Tissue Factor. Amidolytic activity as  $K_M$  [mM],  $k_{cat}$  [s<sup>-1</sup>] and percentage  $k_{cat}/K_M$  compared to FVIIa, pABA inhibition as percentage  $K_i$  of FVIIa and Factor X activation as percentage of FVIIa. More active or stronger binding variants are shown in blue and less active or weaker binding variants in red. All data is shown as the mean of duplicate runs with 3  $\mu$ M soluble tissue factor. Variants are sorted by loop length and deletion position (in red) with the start of each sliding window underlined and the resulting loop length listed.



**Supplemental Figure S3. (A)** Overview of the reconstructed 170-loop for all 39 variants ordered according to loop length and deletion position using FVIIa-WT (PDB ID: 1dan) as a template. High variation in the position traced by the Rosetta NGK protocol is shown in red, low is shown in blue. (B) Only positions 311{c169}-321{c170I} was rebuild using the protocol as shown on the template structure of FVIIa-WT.



**Supplemental Figure S4.** Overview of model energy (in Rosetta Energy Units; REU) calculated using the Talaris2014 energy function plotted against root-mean-square deviation for all 39 generated variants with the FVIIa-WT structure (PDB ID: 1dan) as a template.



**Supplemental Figure S5.** (A) Overview of the reconstructed 170-loop for all 39 variants ordered according to loop length and deletion position using FVIIa- $Y_T$  (PDB ID: 4z6a) as a template. High variation in the position traced by the Rosetta NGK protocol is shown in red, low is shown in blue. (B) Only position  $311\{c169\}-321\{c170I\}$  was rebuild using the protocol as shown on the template structure of FVIIa- $Y_T$ .



**Supplemental Figure S6.** Overview of model energy (in Rosetta Energy Units; REU) calculated using the Talaris2014 energy function plotted against root-mean-square deviation for all 39 generated variants with the FVIIa- $Y_T$  structure (PDB ID: 4z6a) as a template.



**Supplemental Figure S7.** Comparison of crystal structures for wild type trypsin (PDB ID: 1j8a) and FVIIa variant no. 36 build on FVIIa PDB ID 4z6a. Inserts show zoomed views of the 170 loop and the S1 substrate pocket, respectively.



Supplemental Figure S8 FX-activation kinetics of selected FVIIa-variants visualized by SDS-PAGE. (A) Initial FXactivation rate estimated by gel densitometric analysis of FXa HC $\beta$  (c-band) product band on gels. (B) SDS-PAGE analysis of FX-activation by selected FVIIa-variants. MW = molecular weight marker, Xa = FXa control. Major protein bands observed on gel: a = FX HC, b = FXa HC $\alpha$ , c = FXa HC $\beta$  and d = FX/FXa LC