Footprinting haemostatic networks

Eduard H. T. M. Ebberink

Cover design by Robin de Kruijff on behalf of Motion Paintings. Illustration is based on the FIXa structure 1RFN published by Hopfner *et al*.

The research described in this thesis was carried out at the Department of Plasma Proteins, Sanquin Amsterdam, the Netherlands.

ISBN 978-94-6233-810-4

Printed by: Gildeprint - The Netherlands

Financial support for printing was kindly provided by Sanquin Research.

Footprinting haemostatic networks

Footprinting van netwerken in de hemostase
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht op gezag van de rector magnificus, prof.
dr. G.J. van der Zwaan, ingevolge het besluit van het college
voor promoties in het openbaar te verdedigen op maandag
11 december 2017 des ochtends te 10.30 uur

door

Eduard Hubert Theodoor Marius Ebberink

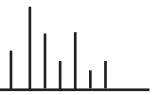
geboren op 07 juni 1985 te Eindhoven

Promotoren: Prof. dr. K. Mertens
Prof. dr. A.B. Meijer

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van de Landsteiner Stichting voor Bloedtransfusieresearch.

Table of contents

1



Introduction

SCOPE OF THIS THESIS

The blood coagulation system involves a delicate interplay between pro-coagulant and anti-coagulant factors that together preserve the flow of blood.[1,2] Proper circulation is threatened by defects of this haemostatic network resulting in either thrombosis or excessive bleeding. Predominantly, the coagulation network comprises enzymes, cofactors and inhibitors that mostly share their structure and functional mechanism. Over the past decades much research has been done to unravel determinants for interactions between these haemostatic proteins, both with each other and with components in their environment. Remarkably, several interactions with cellular components are rather unspecific such as the binding of coagulation factors IX (FIX) and VIII (FVIII) to the promiscuous LDL receptor-related protein-1 (LRP).[3-6] On the other hand, reciprocal interactions between haemostatic proteins are highly specific such as complex assembly of activated FIX (FIXa) and FVIII (FVIIIa) on a phospholipid membrane.[7] Despite the overall structural homology within the family of coagulation factors, they are subject to subtle regulation by a diversity of mechanisms. This not only implies a fine line between specificity and promiscuity, but also introduces elusive allosteric responses in the coagulation system. To date, such paradoxes within the field of haemostasis remain puzzling.

Until recently, the emphasis has been on defining structural and functional determinants on coagulation proteins in well-defined in vitro model systems. Crystallography has provided insight into the 3-dimensional structure of a variety of isolated haemostatic proteins. While these studies have provided us with a detailed structural view on the molecules involved, the inherent rigidity of crystal structures frequently limits the translation of structure into to the biological function. Insight into protein dynamics is needed in order to really understand the catalytic events and assembly of protein complexes within the coagulation cascade. In the present thesis, this issue has been addressed by recent innovative functional proteomics approaches such as chemical labelling using isobaric tandem-mass-tags (TMTs) and hydrogen-deuterium exchange (HDX). Using these mass spectrometry-based techniques attempts have been made to monitor changes in dynamics and conformation of haemostatic proteins. Chemical "footprinting" by TMTs and HDX allows for the identification of exposed surface elements that provide a clue for their interactions or for conformational changes upon ligand/substrate binding, activation and domain swapping.

Validation of footprinting by TMTs has been performed by studying the interaction between LRP and its natural antagonist receptor-associated-protein RAP (**Chapter 2**). This interaction is known to involve several critical lysine residues of RAP that interact with negatively charged patches in LRP. It has been shown that

coagulation factors interact with LRP in a similar manner.[8] In a more explorative setting, TMTs have been used to address the protein integrity and interdomain interaction of two FVIII variants that bear the C-domains of its homologous cofactor FV (**Chapter 3**). These FVIII/FV C-domain chimeras were used to probe determinants for FVIII-specific properties such as interaction with FIXa and cellular targeting to designated von Willebrand Factor carrying organelles (Weibel-Palade bodies) (**Chapter 3**). Mass spectrometry footprinting was combined with kinetic studies and computer modelling to unravel the mechanism behind the pathogenic polymorphism G221E (Marburg-1) in the plasma serine protease Hyaluronan-Binding Protein 2 (HABP2) (**Chapter 4**). Although the role of HABP2 in blood coagulation remains elusive, we found that Gly-221 on the protease domain 220-loop has a vital part in its catalytic function. This is compatible with the concept that protease domain surface loops contribute to the protease activity of coagulation factors.[9] To unbiasedly explore this view, coagulation factors IX, X and prothrombin have been subjected to TMT footprinting to identify differential mechanisms in the protease domain and its surface loops (**Chapter 5**). Special focus had been given to FIX and its 220-loop (**Chapter 6**). We particularly addressed the notion that, following activation, FIXa exhibits remarkably poor enzymatic activity compared to the other coagulation factors. Using TMT footprinting and HDX techniques we identified some intramolecular interactions that regulate FIXa activity.

ALLOSTERY IN THE COAGULATION SYSTEM

The fine line between promiscuity and specificity is particularly visible amongst the serine proteases of the coagulation system. Coagulation serine proteases such as thrombin, activated FIX (FIXa), FVII (FVIIa) and FX (FXa) share their proteolytic active domain with that of the chymotrypsin superfamily and are highly homologous (Fig. 1). While the protease domain architecture is essentially identical, they are selectively activated and operate with high specificity towards their substrates and cofactors. The coagulation factors appear particularly distinct with regard to their response to ligands that stimulate or inhibit their activity.

Compared to the parent protease chymotrypsin, thrombin has two distinctive properties: i) the protease domain requires Na^+ to become fully active and ii) depending on the ligand, thrombin changes affinity for its substrate and can become either pro-coagulant (for instance due to Na^+-binding) or anti-coagulant (due to binding to the endothelial membrane protein thrombomodulin).[10-12] Especially for Na^+-dependency, the allosteric mechanism has extensively been studied. Thereby, in light

of recent interpretations on allostery, different theories have been proposed. In one theory, Na^+-binding is thought to be one of the key components in stabilization of some protease domain surface loops.[13-15] The reduction in protein flexibility produces a 'continuum' of shifting conformational states. Along this thermodynamic trajectory, thrombin can be shuttled between zymogen- and proteinase-like states.[16,17] In another hypothesis, the emphasis lays on the 215-217 β-strand that flanks the protease domain active site. Ligands such as Na^+ could allosterically drive the strand onto the active site or away from it. Thereby, the protease domain switches between an open (E) or closed conformation (E*).[18] Recently, on the basis of rapid kinetics, distinct conformations (e.g. E or E*) have been inferred to result from ligand binding (selection of conformation) for thrombin and other trypsin-like proteases.[19]

Similar allosteric mechanisms have been proposed for coagulation factor X.[20-22] Also FX binds Na^+ with concomitant enhancement in catalytic activity.[23,24] However, FXa requires interaction with additional factors. For instance, the protease domain also requires Ca^{2+} which binds on a separate location opposite the Na^+-binding site (Fig. 1). Furthermore, activity is enhanced by assembly into the prothrombinase complex with coagulation factor V (FV) on a phospholipid membrane.[25] The exact mechanism behind the Ca^{2+}- and FV-driven enhancement remains unresolved. Current understanding places these allosteric effects in line with previous mentioned hypotheses.[21,26] An additional aspect in FX is that, following activation, the light chain remains attached and has an essential contribution to activity and complex assembly.[27] Ca^{2+} is also required for the FXa light chain to drive complex assembly on negatively charged lipids.

Activated coagulation factor IX highly resembles FXa (Fig. 1). Both its heavy chain and its light chain require Ca^{2+}. Similar to FXa, FIXa exhibits Na^+-dependence too.[28] However, the effect is less evident compared to other coagulation factors, especially when Ca^{2+} is present.[29] The cofactor of FIXa, FVIII, displays high homology to FV and participates in membrane-bound complex that is almost identical to the prothrombinase complex. These events are able to 'switch-on' the catalytic activity of FIXa which on its own is a very poor enzyme. Previous research from our group and others focused on sites in the light chain subunits that translate into complex assembly and/or protease domain activity.[30,31] Several sites in the FIXa light chain have been reported to be important in interdomain cross-talk between the light and heavy chain.[32]

How allosteric interactions translate into the dramatic impact on catalytic function of thrombin, FXa and FIXa remains poorly understood. It seems evident that this should involve some intramolecular signalling. In many of these questions, structural techniques (X-ray crystallography, Electron Microscopy, etc.) have provided much insight in conformational communication. Unfortunately, the 'structure-focused'

approach does not provide a full explanation for the complex, thermodynamically linked events.[33-35] Inspection of static, 'end point' protein structures ignores the intrinsic dynamic nature of proteins.[33,36] Taking advantage of the current knowledge on crystal structures, we now aimed to expand our insight into the allosteric mechanism by use of novel mass spectrometry techniques.

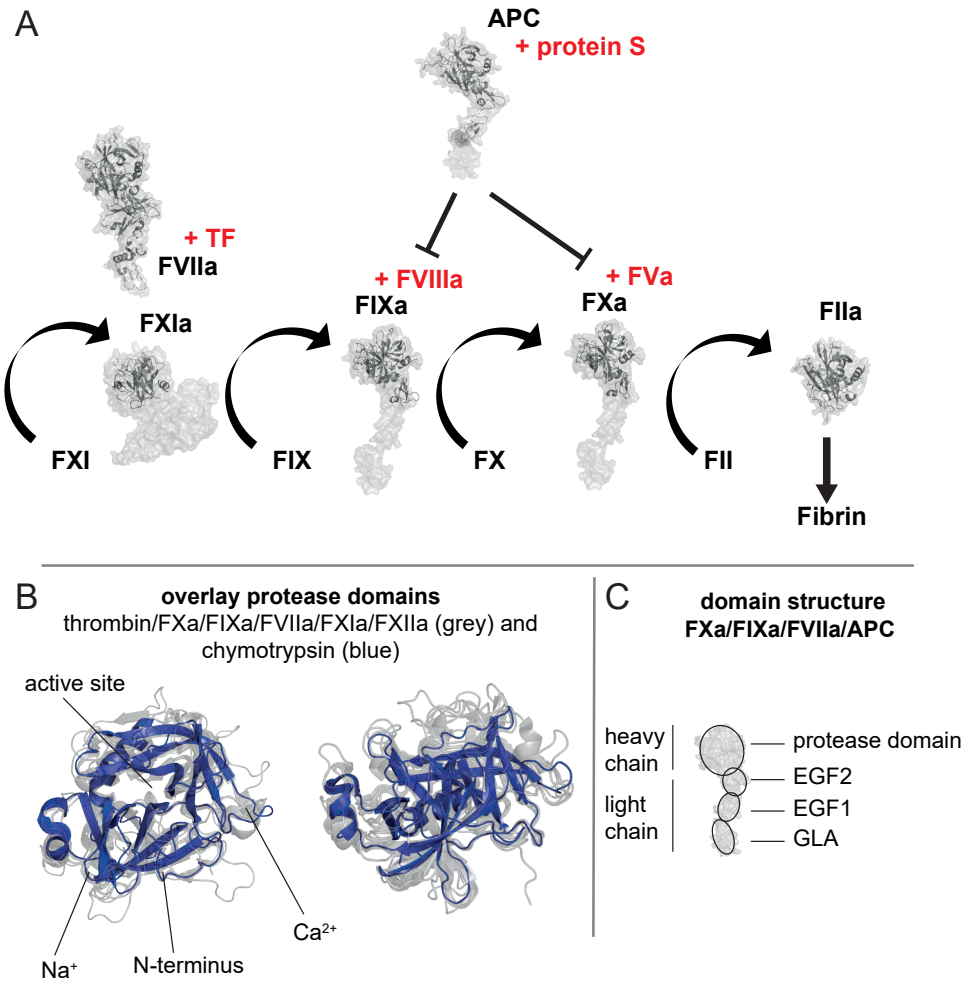


Figure 1 *(figure legend on next page)*

Figure 1: **Homology in the coagulation cascade. (A)** Part of the coagulation cascade involving serine proteases FXI, FIX, FX, activated protein C (APC) and thrombin. These proteases share much of their domain structure. In particular, the catalytic domains are shared among coagulation serine proteases. Crystallographic structures are displayed as grey cartoons. Proteases of the coagulation cascade are activated by limited proteolysis. Cleavage of FIX can be catalysed by FXIa as well as by FVIIa. For full catalytic activity, however, FVIIa requires its cofactor: tissue factor (TF). On its turn, FIXa is able to convert FX into FXa. Also this process requires the interaction with a cofactor (FVIIIa). Activated FX is able to activate prothrombin and also needs a cofactor to achieve optimal catalytic activity. APC, a serine protease with homology to FIXa and FXa, inhibits the coagulation process by cleavage of cofactors FVIIIa and FVa. As for these cofactors, both FVIIIa and FVa share a homologous domain organization. Protein S functions as cofactor for APC. **(B)** To visualize the homology in protease domains, X-ray crystallography structures of the protease domain of different coagulation factors (PDB codes: thrombin:1PPB, FXa:2BOK, FIXa:1RFN, APC:1AUT, FVIIa:1DAN, FXIa:3SOS, FXIIa:4XE4) are superimposed on that of chymotrypsin (1GCT, coloured blue). Top and side views are given. The active site, N-terminus insertion site, Ca^{2+} and Na^{+} binding sites are indicated. **(C)** Besides the protease domain, several coagulation factors share the same domain topology. For instance, FIXa, FXa, FVIIa and APC have a similar light chain comprised of EGF1, EGF2 and Gla domains.

PROBING CONFORMATIONAL CHANGE BY MASS SPECTROMETRY

Mass spectrometry (MS) monitors the path of a charged particle that is placed in an electric or magnetic field to accurately determine its mass-to-charge ratio. Precise mass-to-charge ratios could be used to derive composition of the assessed particle, whether or not with subsequent fragmentations of the particle. With the development of soft ionization techniques (electron spray ionization and matrix assisted laser desorption ionization) mass spectrometry gained much attention in structural biology. For protein research this primarily entails identification of the peptide primary sequence by measuring masses of peptide fragments. By virtue of these advances, MS has been increasingly used to detect and quantify proteins, protein-protein complexes and post-translational modifications.[37-39]

Besides protein identification and quantification, MS could provide information on the structure of proteins after chemical modification of amino acid residues (with corresponding shift in mass) prior to MS analysis. Thereby, the protein architecture is probed by the propensity of amino acid residues to become modified in a chemical 'footprinting' approach. Footprinting infers that the more exposed regions in the protein are easier modified since they are more accessible. Chemical crosslinking or surface labelling of proteins in conjunction with mass spectrometry has proven to be powerful tools in exploring change of protein conformation and

surface exposure.[40,41] An advantageous aspect thereof is that the chemical footprinting reaction is performed in an aqueous environment. Upon in-solution labelling, the physiological conditions are preserved and the dynamic nature of proteins is not disrupted. Therefore, following MS analysis, structural information is obtained on an atomic level while dynamics are maintained. Over the years, various types of footprinting techniques have been developed to assess multiple aspects of protein architecture and dynamics.

Footprinting by selective, covalent labelling

An established method of footprinting is the chemo-selective labelling of amino acid side chains. Residues that are readily modified under (near) physiological conditions are charged residues Asp, Glu, Cys, His, Arg and Lys.[42] Also, under more stringent conditions aromatic residues Trp and Tyr could be targeted for labelling. Covalent, irreversible labelling techniques have the advantage that modified sites are more robust towards downstream processing for subsequent mass spectrometry analysis. Furthermore, functionality of the tag itself could be taken into consideration. For instance, isotopically labelled or biotinylated tags have been developed to respectively facilitate quantification and enrichment of labelled peptides.[38,43] In the footprinting study of HABP2, we have employed a biotin-carrying tag to facilitate enrichment and identification of one specific site in the protein, the N-terminus of the protease domain (**Chapter 4**).

In the studies presented in this thesis, we make use of isobaric tandem-mass-tags (TMTs) to footprint plasma proteins and explore their structural behaviour. By isobaric mass tags, ambiguities due to protein concentration and side chain modification can be accounted for by comparative labelling of a protein under two different conditions. Besides a reactive group, TMTs contain special reporter groups with differentially distributed isotopes (TMT-126 and TMT-127). As side chain modification, both tags maintain the same molecular mass (Fig. 2A). Only upon fragmentation in the mass spectrometer the reporter group is disintegrated and different isotope distributions are detected with release of an indicator mass (Fig. 2B). The extent of labelling of a residue is given by the indicator masses (126 Da and 127 Da) in the MS^2 spectra of the peptide (Fig. 2B). Due to the isobaric nature of TMTs, proteins can be labelled by essentially the same tag under different conditions. Downstream processing such as denaturation and proteolysis is identical for the labelled samples. Quantification is done simultaneously as the identification of the labelled site (Fig. 2B). Two kinds of TMTs have been developed with different reactive groups: one with an iodoacetyl group reactive towards cysteine residues and one containing an N-hydroxysuccinimide (NHS) group reactive towards the primary amine of lysine residues (Fig. 2A). The abundance of exposed free cysteine residues in

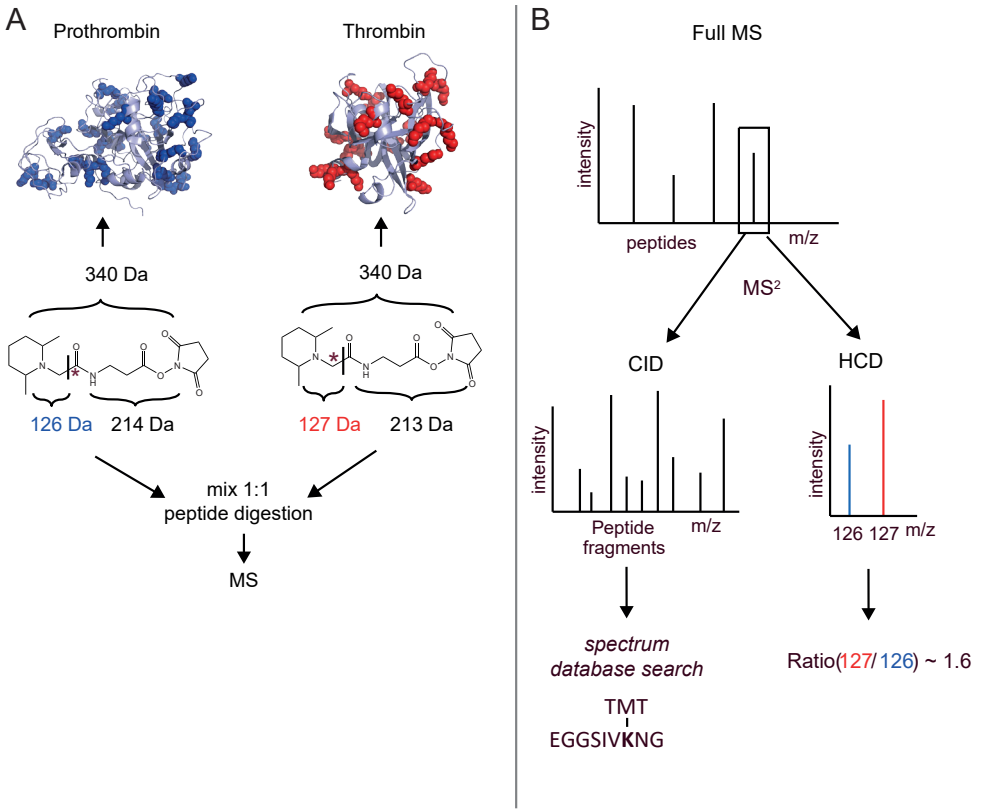


Figure 2: **Mass spectrometry footprinting by tandem-mass-tags (TMTs). (A)** As an example, lysine residues of prothrombin (PDB code 4HZH) and thrombin (PDB code 1PPB) are labelled using the two different types of TMTs (TMT-126 and TMT-127 respectively). Overall both TMTs have the same molecular mass (340 Da) and also the modification mass is the same (+225 Da). The difference in TMTs is present in distribution of the isotopes. This becomes visible only upon tandem mass spectrometry. Due to the fragmentation a reporter group can be released with either has a mass of 126 or 127 Da. Labelled prothrombin and thrombin are mixed with equimolar amounts and processed into peptides. Peptides are subjected to tandem mass spectrometry to quantify the amount of labelling. By doing so, the extent of labelling of the same lysine residue in prothrombin and thrombin can be compared. **(B)** A schematic overview of the data-dependent acquisition in the mass spectrometer which was done to identify and quantify labelled lysine residues. Initially, the mass spectrometer performs a full scan which monitors the influx of injected peptides. The spectrum indicates different m/z intensities that potentially represent labelled peptides. Peaks that are detected with high intensities are selected for further fragmentation. With collision induced dissociation (CID) fragmentation and a subsequent database search, the composition of labelled peptides can be determined. During the full scan and CID fragmentation, the TMT-moiety on the lysine residue remains intact. Only upon higher energy collisional dissociation (HCD) fragmentation the reporter group is released. Therefore, directly following the CID fragmentation, HCD is applied on the same peptide. Intensities of the TMT reporter groups then can be used to quantify labelling of the lysine residue. This is expressed as a ratio of TMT-127/TMT-126.

nature is limited (1.1%) due to their high reactivity.[44] Therefore, we used NHS reactive TMTs to examine accessibility of the lysine residues because of their high abundancy in proteins and protein surface.

Initially, the TMT-labelling approach was tested on the LRP-RAP protein complex which is known to involve certain critical lysine residues.[45] (**Chapter 2**). To assess protein integrity and differences in conformation, the TMT-labelling of mutant and wild-type FIX and FVIII was compared (**Chapter 3 and 6**). Moreover, we expanded the TMT-labelling approach in a time-resolved fashion to probe difference between activation of FIX, FX and prothrombin (**Chapter 5**).

Footprinting by hydrogen/deuterium exchange

Compared to previously discussed covalent labelling techniques, hydrogen/deuterium exchange (HDX) represents a more reversible labelling method. In a strict sense, HDX also involves covalent modification, namely of the protein O-H, N-H and S-H bonds. These bonds appear rather labile and allow for the exchange of hydrogens with surrounding water. This phenomenon is utilized by incubation of protein in D_2O. Exchange of hydrogens with deuterium and subsequent incorporation of deuterium in the protein can readily be detected in the mass spectrometer. Essential is that exchange rates are influenced by several biochemical features. For instance, the hydrogens in C-H bonds display negligible exchange.[46] On the other hand, hydrogens of amino acid side chains exchange too fast to be measured. In fact, only backbone amide hydrogens are considered measurable with exchange rates ranging from seconds to hours and beyond at physiological pH.[47]

The element that makes HDX suitable as footprinting technique is that exchange rates are dictated, alongside pH and temperature, through protein conformation. Conformational changes can decrease exchange rates by a factor of 10^8. Within the protein structure, especially hydrogen bonds are considered critical because they essentially block hydrogens from exchanging even when fully accessible to the solvent.[48] Therefore, structural motifs such as α-helices and β-sheets which contain many hydrogen bonds are considered relatively protected from HDX. Hydrogen bonded structures must open up first to be able to exchange with the environment.

To assess these HDX kinetics, we performed continuous labelling experiments. This entails dilution of the protein in deuterium and subsampling over time to determine the uptake of deuterium. The exchange rates measured at various time-points gives information on protein dynamics in its native state and under equilibrium conditions. Here, we employed continuous labelling on FIX and FIXa to probe differences in the conformational dynamics (**Chapter 6)**. Besides the natural 'molecular breathing', local folding/unfolding events could also allosterically be induced by ligand-binding. This

possibility is explored for FIXa binding a synthetic pseudo-substrate in the active site (FIXa-EGR) (**Chapter 6**). Finally, the merit of these novel footprinting methods is discussed in the last chapter (**Chapter 7**).

REFERENCES

1. Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry.* 1991 **30** 10363-10370.
2. Mann KG, Nesheim ME, Church WR, Haley P, Krishnaswamy S. Surface-dependent reactions of the vitamin K-dependent enzyme complexes. *Blood.* 1990 **76** 1-16.
3. Neels JG, van Den Berg BM, Mertens K, ter Maat H, Pannekoek H, van Zonneveld AJ, Lenting PJ. Activation of factor IX zymogen results in exposure of a binding site for low-density lipoprotein receptor-related protein. *Blood.* 2000 **96** 3459-3465.
4. Kurasawa JH, Shestopal SA, Karnaukhova E, Struble EB, Lee TK, Sarafanov AG. Mapping the binding region on the low density lipoprotein receptor for blood coagulation factor VIII. *J Biol Chem.* 2013 **288** 22033-22041.
5. Kurasawa JH, Shestopal SA, Woodle SA, Ovanesov MV, Lee TK, Sarafanov AG. Cluster III of low-density lipoprotein receptor-related protein 1 binds activated blood coagulation factor VIII. *Biochemistry.* 2015 **54** 481-489.
6. Rohlena J, Kolkman JA, Boertjes RC, Mertens K, Lenting PJ. Residues Phe342-Asn346 of activated coagulation factor IX contribute to the interaction with low density lipoprotein receptor-related protein. *J Biol Chem.* 2003 **278** 9394-9401.
7. Mertens K, Celie PH, Kolkman JA, Lenting PJ. Factor VIII-factor IX interactions: molecular sites involved in enzyme-cofactor complex assembly. *Thromb Haemost.* 1999 **82** 209-217.
8. Van den Biggelaar M, Madsen JJ, Faber JH, Zuurveld MG, van der Zwaan C, Olsen OH, Stennicke HR, Mertens K, Meijer AB. Factor VIII Interacts with the Endocytic Receptor Low-density Lipoprotein Receptor-related Protein 1 via an Extended Surface Comprising "Hot-Spot" Lysine Residues. *J Biol Chem.* 2015 **290** 16463-16476.
9. Huntington JA. Thrombin plasticity. *Biochim Biophys Acta.* 2012 **1824** 246-252.
10. Orthner CL, Kosow DP. Evidence that human alpha-thrombin is a monovalent cation-activated enzyme. *Arch Biochem Biophys.* 1980 **202** 63-75.
11. Dang OD, Vindigni A, Di Cera E. An allosteric switch controls the procoagulant and anticoagulant activities of thrombin. *Proc Natl Acad Sci U S A.* 1995 **92** 5977-5981.
12. Bode W. The structure of thrombin: a janus-headed proteinase. *Semin Thromb Hemost.* 2006 32 Suppl 1 16-31.
13. De Filippis V, De Dea E, Lucatello F, Frasson R. Effect of Na^+ binding on the conformation, stability and molecular recognition properties of thrombin. *Biochem J.* 2005 **390** 485-492.

14. Parry MA, Stone SR, Hofsteenge J, Jackman MP. Evidence for common structural changes in thrombin induced by active-site or exosite binding. *Biochem J.* 1993 **290** 665-670.
15. Lechtenberg BC, Johnson DJ, Freund SM, Huntington JA. NMR resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc Natl Acad Sci U S A*. 2010 **107** 14087-14092.
16. Kamath P, Huntington JA, Krishnaswamy S. Ligand binding shuttles thrombin along a continuum of zymogen- and proteinase-like states. *J Biol Chem*. 2010 **285** 28651-28658.
17. Lechtenberg BC, Freund SM, Huntington JA. An ensemble view of thrombin allostery. *Biol Chem.* 2012 **393** 889-898.
18. Di Cera E. Thrombin. *Mol Aspects Med.* 2008 **29** 203-254.
19. Vogt AD, Pozzi N, Chen Z, Di Cera E. Essential role of conformational selection in ligand binding. *Biophys Chem*. 2014 **186** 13-21.
20. Toso R, Zhu H, Camire RM. The conformational switch from the factor X zymogen to protease state mediates exosite expression and prothrombinase assembly. *J Biol Chem.* 2008 **283** 18627-18635.
21. Underwood MC, Zhong D, Mathur A, Heyduk T, Bajaj SP. Thermodynamic linkage between the S1 site, the Na^+ site, and the Ca^{2+} site in the protease domain of human coagulation factor Xa. Studies on catalytic efficiency and inhibitor binding. *J Biol Chem.* 2000 **275** 36876-36884.
22. Gohara DW, Di Cera E. Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends Biotechnol.* 2011 **29** 577-585.
23. Orthner CL, Kosow DP. The effect of metal ions on the amidolytic acitivity of human factor Xa (activated Stuart-Prower factor). *Arch Biochem Biophys.* 1978 **185** 400-406.
24. Schärer K, Morgenthaler M, Paulini R, Obst-Sander U, Banner DW, Schlatter D, Benz J, Stihle M, Diederich F. Quantification of cation-pi interactions in protein-ligand complexes: crystal-structure analysis of Factor Xa bound to a quaternary ammonium ion ligand. *Angew Chem Int Ed Engl*. 2005 **44** 4400-4404.
25. Rezaie AR, He X. Sodium binding site of factor Xa: role of sodium in the prothrombinase complex. *Biochemistry.* 2000 **39** 1817-1825.
26. Camire RM. Prothrombinase assembly and S1 site occupation restore the catalytic activity of FXa impaired by mutation at the sodium-binding site. *J Biol Chem.* 2002 **277** 37863-37870.
27. Wilkinson FH, Ahmad SS, Walsh PN. The factor IXa second epidermal growth factor (EGF2) domain mediates platelet binding and assembly of the factor X activating complex. *J Biol Chem.* 2002 **277** 5734-5741.
28. Schmidt AE, Stewart JE, Mathur A, Krishnaswamy S, Bajaj SP. Na^+ site in blood

coagulation factor IXa: effect on catalysis and factor VIIIa binding. *J Mol Biol.* 2005 **350** 78-91.
29. Gopalakrishna K, Rezaie AR. The influence of sodium ion binding on factor IXa activity. *Thromb Haemost.* 2006 **95** 936-941.
30. Celie PH, Van Stempvoort G, Fribourg C, Schurgers LJ, Lenting PJ, Mertens K. The connecting segment between both epidermal growth factor-like domains in blood coagulation factor IX contributes to stimulation by factor VIIIa and its isolated A2 domain. *J Biol Chem*. 2002 **277** 20214-20220.
31. Yang X, Chang YJ, Lin SW, Walsh PN. Identification of residues Asn89, Ile90, and Val107 of the factor IXa second epidermal growth factor domain that are essential for the assembly of the factor X-activating complex on activated platelets. *J Biol Chem*. 2004 **279** 46400-46405.
32. Fribourg C, Meijer AB, Mertens K. The interface between the EGF2 domain and the protease domain in blood coagulation factor IX contributes to factor VIII binding and factor X activation. *Biochemistry*. 2006 **45** 10777-10785.
33. Motlagh HN, Wrabl JO, Li J, Hilser VJ. The ensemble nature of allostery. *Nature*. 2014 **508** 331-339.
34. Liu J, Nussinov R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLoS Comput Biol*. 2016 **12** e1004966.
35. Cui Q, Karplus M. Allostery and cooperativity revisited. *Protein Sci*. 2008 **17** 1295-1307.
36. Cooper A, Dryden DT. Allostery without conformational change. A plausible model. *Eur Biophys J*. 1984 **11** 103-109.
37. Lössl P, van de Waterbeemd M, Heck AJ. The diverse and expanding role of mass spectrometry in structural and molecular biology. *EMBO J*. 2016 **35** 2634-2657.
38. Kovanich D, Cappadona S, Raijmakers R, Mohammed S, Scholten A, Heck AJ. Applications of stable isotope dimethyl labeling in quantitative proteomics. *Anal Bioanal Chem*. 2012 **404** 991-1009.
39. Yang Y, Franc V, Heck AJR. Glycoproteomics: A Balance between High-Throughput and In-Depth Analysis. *Trends Biotechnol*. 2017 **35** 598-609.
40. Sinz A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom Rev.* 2006 **25** 663-682.
41. Petrotchenko EV, Serpa JJ, Borchers CH. Use of a combination of isotopically coded cross-linkers and isotopically coded N-terminal modification reagents for selective identification of inter-peptide crosslinks. *Anal Chem*. 2010 **82** 817-823.
42. Mendoza VL, Vachet RW. Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev*. 2009 **28** 785-815.
43. Ori A, Free P, Courty J, Wilkinson MC, Fernig DG. Identification of heparin-binding

sites in proteins by selective labeling. *Mol Cell Proteomics.* 2009 **8** 2256-2265.

44. Marino SM, Gladyshev VN. Cysteine function governs its conservation and degeneration and restricts its utilization on protein surfaces. *J Mol Biol*. 2010 **404** 902-916.
45. Fisher C, Beglova N, Blacklow SC. Structure of an LDLR-RAP complex reveals a general mode for ligand recognition by lipoprotein receptors. *Mol Cell*. 2006 **22** 277-283.
46. Gallagher ES, Hudgens JW. Mapping Protein-Ligand Interactions with Proteolytic Fragmentation, Hydrogen/Deuterium Exchange-Mass Spectrometry. *Methods Enzymol.* 2016 **566** 357-404.
47. Morgan CR, Engen JR. Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr Protoc Protein Sci.* 2009 Chapter 17:Unit 17.6.1-17.
48. Skinner JJ, Lim WK, Bédard S, Black BE, Englander SW. Protein dynamics viewed by hydrogen exchange. *Protein Sci.* 2012 **21** 996-1005.

2

A novel chemical footprinting approach identifies critical lysine residues involved in the binding of receptor-associated protein to cluster II of LDL receptor-related protein

Esther Bloem[1*], Eduard H.T.M. Ebberink[1*], Maartje van den Biggelaar[1,2], Carmen van der Zwaan[1], Koen Mertens[1] and Alexander B. Meijer[1,2]

Biochem J. 2015 468(1):65-72

[1]Department of Plasma Proteins, Sanquin Research, 1066 CX Amsterdam, The Netherlands
[2]Pharmaceutical Sciences, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands

**authors contributed equally*

ABSTRACT

Tandem-mass-tags (TMTs) were utilized in a novel chemical footprinting approach to identify lysine residues that mediate the interaction of receptor-associated-protein (RAP) with cluster II of LDL receptor-related protein. The isolated RAP D3 domain was modified with TMT-126 and the D3 domain-cluster II complex with TMT-127. Nano-LC mass spectrometry analysis revealed reduced modification with TMT-127 of peptides including Lys256, Lys270 and Lys305-Lys306 suggesting that these residues contribute to cluster II binding. This agrees with previous findings that Lys256 and Lys270 are critical for binding cluster II sub-domains [Fisher, Beglova and Blacklow (2006) Mol. Cell 22, 277-283]. Cluster II-binding studies utilizing D3 domain variants Lys256Ala, Lys305Ala, and Lys306Ala now showed that Lys306 contributes to cluster II binding as well. For full length RAP, we observed that peptides including Lys60, Lys191, Lys256, Lys270, and Lys305-Lys306 exhibited reduced modification with TMT in the RAP-cluster II complex. Notably, Lys60 has previously been implicated to mediate D1 domain interaction with cluster II. Our results suggest that also Lys191 of the D2 domain contributes to cluster II binding. Binding studies employing the RAP variants Lys191Ala, Lys256Ala, Lys305Ala and Lys306Ala, however, revealed a modest reduction in cluster II binding for the Lys256Ala variant only. This suggests that the other lysine residues can compensate for the absence of a single lysine residue for effective complex assembly. Collectively, novel insight has been obtained about the contribution of lysine residues of RAP to cluster II binding. In addition, we propose that TMTs can be utilized to identify lysine residues critical for protein complex formation.

Summary statement

Mass spectrometry was employed to explore receptor interactive regions on receptor-associated-protein. A footprinting approach identified established and novel lysine residues contributing to receptor binding. The developed approach can be utilized to screen protein complexes for critical interactive lysine residues.

INTRODUCTION

The assembly of a protein complex is a key event in almost all biological processes. To gain insight into these mechanisms and defects therein, it is of critical importance to understand how proteins interact. Yet, identification of protein interaction sites has remained a continuous challenge.[1] An evolving technology for the identification of interactive regions on proteins involves chemical modification of amino acid regions of unbound and bound proteins.[2-4] The amino acid residues that contribute to complex assembly are expected to be protected from chemical modification. In 2009, Ori *et al*. demonstrated in an elegant study that heparin binding sites can be identified employing a selective labeling strategy.[5]

A major issue to overcome using a chemical footprinting approach is that the equilibrium between bound and unbound states of proteins will also allow for unintended chemical modification of residues that are critical for the interaction. Especially for proteins that bind with low affinity, the specific protection from chemical modification may be hardly, or not at all, observed. To overcome this issue, footprinting approaches have been developed that rely on fast modification of protein complexes. These methods include, for instance, hydroxyl radical oxidation of amino acid regions of a protein complex.[3]

Previously, we and others employed lysine-directed isobaric tandem-mass-tags to assist in the structural characterization of proteins using mass spectrometry.[6-8] In the present study, we explore the potential of these mass tags to effectively identify lysine residues that directly contribute to protein complex formation. In the employed approach, we make use of the isobaric tandem-mass-tags TMT-126 and TMT-127. Modification of the lysine residues of assembled proteins with TMT-127 and the unbound proteins with TMT-126 is expected to lead to a reduced incorporation with TMT-127 in amino acid regions comprising lysine residues that contribute to complex formation. Taking maximum advantage of the fact that TMT-126 and TMT-127 exhibit a different isotope distribution, MS/MS fragmentation of modified peptides derived from these regions allows for relative quantification of the incorporation with TMT-126 and TMT-127.[9] As a model system, we employ the complex between receptor-associated-protein (RAP) and cluster II of the ligand-binding domains of LDL receptor-related protein (LRP). This complex is particularly suitable for this study as (i lysine residues have been shown to contribute to complex formation,[10-12] (ii the identity of three critical lysine residues of RAP has been established from crystallography and NMR studies,[10,13] and (iii the contribution of additional lysine residues remains to be assessed.

LRP is a member of the LDL receptor (LDLR) family, and has been implicated to play a role in a range of biological processes including cell migration, vascular

permeability, and the catabolism of coagulation proteins.[14] The physiological role of RAP is to assist in the proper intracellular folding of the LDLR family members, and to prevent premature intracellular ligand-binding.[15] Because of the latter characteristic, RAP has been frequently employed as an antagonist to identify novel binding partners of LRP.[16] In addition, the complex between RAP and ligand-binding domains has been studied to gain insight into the general mechanism of complex assembly between LRP and its ligands.[11-13,15-20]

RAP comprises three similar D domains each of which has been demonstrated to bind LDLR and LRP. Although there is a debate in literature about the actual binding affinities of the individual domains, it has been proposed that the isolated D3 domain binds more effective to the LRP/LDLR ligand-binding domains than the isolated D1 and D2 domains.[18,20] The ligand-binding domains of the LDLR-like proteins are, in turn, small compact domains that are clustered in distinct regions within the protein. LDLR comprises a single cluster containing 7 of these so-called complement type repeats, whereas LRP contains four of these specialized ligand-binding regions.[21]

Crystal structure analysis of the RAP D3 domain in complex with two repeats from LDLR has revealed that Lys256 and Lys270 of the D3 domain are critical for the interaction.[10] The structure shows that each of these lysine residues is inserted into an "acidic necklace" of negatively charged residues of a single complement-type repeat.[10,17] We and others have demonstrated that an arginine residue cannot replace the lysine residue in this binding mechanism.[10,12,13,22] NMR analysis of the RAP D1 domain in interaction with two complement-type repeats from LRP revealed that Lys60 is critical for the interaction with a complement-type repeat.[13] No information is available about the lysine residues of the D2 domain that may interact with an acidic necklace of a ligand-binding domain of LRP.

Using our footprinting approach, we confirm that Lys60, Lys256, and Lys270 contribute to the binding of RAP to LRP cluster II. Our results together further suggest that lysine residues 191 of the D2 domain and 306 of the D3 domain are involved in LRP cluster II binding as well.

EXPERIMENTAL

Proteins

Human RAP D3 domain was purified from DH5α cells as described.[12,19] Full length rat glutathione-S-transferase RAP was expressed and purified as described.[12,16] Variants of human RAP D3 domain and full length rat RAP were constructed using Quick-change mutagenesis (Stratagene, La Jolla, CA, USA) according to the instructions of the manufacturer using appropriate primers. LRP1 cluster II was expressed in Baby Hamster Kidney cells and purified as described.[23] This cluster II fragment contains an amino acid tag which is utilized for detection of cluster II with peroxidase labeled monoclonal antibody CLB-CAg69.[24,25]

TMT modification

Human RAP D3 domain or full length RAP was incubated in presence or absence of a 1- or 10-fold molar excess of LRP cluster II for 15 minutes at 37 °C in 50 mM HEPES pH 7.4, 150 mM NaCl, 5 mM $CaCl_2$. RAP D3 domain or full length RAP (2 µg in total) were incubated with a 10.000-fold molar excess of TMT-126 and the RAP-cluster II complex with TMT-127 for 15 minutes at 37 °C. The TMT-labeling reaction was terminated by the addition of 150-fold molar excess of hydroxylamine over the TMTs. Protein mixtures were pooled at a one-to-one ratio, and the cysteines were alkylated as described.[6] Proteins were proteolysed by either chymotrypsin, Glu-C or Asp-N according to the instructions of the manufacturer (Thermo Fisher Scientific Inc, Bremen, Germany). Obtained peptides were desalted employing a C18 Ziptip (Millipore Corporation, Billerica, USA) according to the instructions of the manufacturer.

Mass spectrometry analysis

Peptides were separated by reverse-phase chromatography and sprayed into a LTQ OrbitrapXL mass spectrometer (Thermo Fisher Scientific Inc, Bremen, Germany) essentially as described.[6,26] During reverse-phase chromatography, we utilized a 40-minute gradient from 0% to 35% (v/v) acetonitrile with 0.5% (v/v) acetic acid. Collision-induced dissociation (CID) spectra and higher energy collision-induced dissociation (HCD) spectra were acquired as described in Dayon *et al*.[9] The three most intense precursor ions in the full scan (300-2000 m/z, resolving power 30.000) with a charge state of 2^+ or higher were selected for CID using an isolation width of 2 Da, a 35% normalized collision energy, and an activation time of 30 ms. The same precursor ions were subjected to HCD with a normalized collision energy of 60%, which allows for the identification of the reporter group from the TMT-label.

Identification of the peptides as well as the TMT-127/TMT-126 ratio thereof

The identification of the peptides and determination of their TMT-127/TMT-126 ratio were assessed employing Proteome Discoverer software 1.2. The SEQUEST search algorithm was used employing the protein database 25.H_sapiens.fasta including the amino acid sequence of human RAP D3 domain or a database comprising RAP from rat. The following selection criteria were used: (i all lysine residues are modified by a TMT-label, (ii all cysteine residues are alkylated, (iii all methionine residues may be oxidized, and (iv a maximum false discovery rate of 5% was accepted. The TMT-ratio of the identified peptides was normalized to the average TMT-ratio obtained within that experiment. We also verified whether the labelling of RAP with TMT-127 and the RAP-cluster II complex with TMT-126 affects the outcome of the experiments (TMT labels are reversed in this experiment). This was, however, not the case (*Supplementary Figure S1*).

Solid phase competition assay

1 µg/ml WT human RAP D3 domain or 0.5 µg/ml full length WT rat RAP was immobilized at 4 °C overnight on a microtiter plate in 0.05 M $NaHCO_3$, pH 9.8. Plates were washed with TBS containing 5 mM $CaCl_2$ and 0.1% tween-20. 2.5 nM LRP cluster II was incubated with RAP D3 domain and 0.5 nM cluster II with full length RAP for 2 hours at 37 °C in the presence of increasing concentrations of RAP D3 domain variants or full length RAP variants. Residual cluster II binding to immobilized RAP D3 domain and full length RAP was detected employing peroxidase labeled monoclonal antibody CAg69 as described.[24,25]

RESULTS

Work flow of the chemical footprinting based mass spectrometry approach using the RAP D3 domain – LRP cluster II complex as model

As crystal structure analysis and mutagenesis studies have demonstrated that the lysine residues 256 and 270 of the D3 domain of RAP bind directly to two complement-type repeats,[10-12] we employed this domain to validate our approach. To this end, the lysine residues of the D3 domain were modified with an excess of TMT-126 in the absence of LRP cluster II, and with TMT-127 in the presence of an excess of LRP cluster II. The modification of the lysine residues was allowed for 15 minutes at 37 °C and the reaction was stopped with hydroxylamine. The proteins were subsequently pooled in equal molar ratio based on the concentration of RAP. Pooled proteins were alkylated and divided into three fractions to enable proteolysis by chymotrypsin,

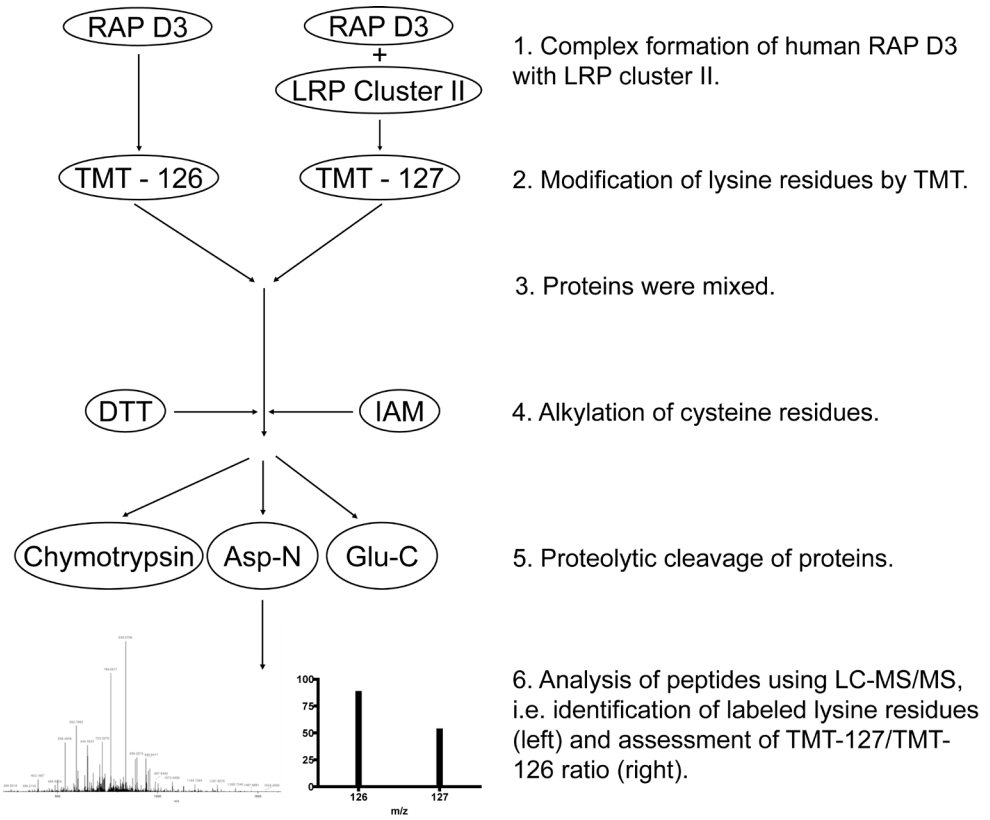


Figure 1: **Work flow of the chemical foot printing approach.** The isolated RAP D3 domain was modified with TMT-126 in the absence of cluster II and with TMT-127 in the presence of cluster II. Proteins were pooled in a one-to-one molar ratio and proteolyzed by chymotrypsin, Asp-N or Glu-C. Peptides were subsequently analyzed on a nano-LC Orbitrap XL mass spectrometer.

Asp-N, and Glu-C. The resulting peptides were analyzed employing a nano-LC LTQ OrbitrapXL mass spectrometer. Collision induced dissociation (CID) of the peptide ions was employed to identify the peptides. Higher Energy CID (HCD) fragmentation of the same peptide was utilized to detect the reporter groups from TMT-126 and TMT-127.[9] Protection of a lysine residue from chemical modification in the presence of LRP cluster II is expected to result in a decreased TMT-127/TMT-126 ratio. A ratio of 1 indicates that the exposure of the involved lysine residue to the solvent is not altered upon complex formation. Figure 1 shows a schematic overview of the work flow.

Identification of lysine residues of the RAP D3 domain that contribute to receptor binding

The peptide mixtures obtained from the above-described approach were analyzed by mass spectrometry. The three most abundant peptide ions between 300 m/z and 2000 m/z in each full scan were subjected to CID and HCD. In total, 23 unique peptides of the D3 domain were identified covering 88% of the complete sequence and all lysine residues of this domain (*Supplementary Table S1*). Figure 2 shows part of the HCD spectra of the TMT-modified peptides 251-EAKIEKHNHY-260, 256-KHNHYQKQLE-265, and 266-IAHEKLRHAE-275. The average TMT-127/TMT-126 ratio obtained from at least four independent experiments is displayed in Figure 3. The data revealed a decrease in the ratio for the peptides including the lysine

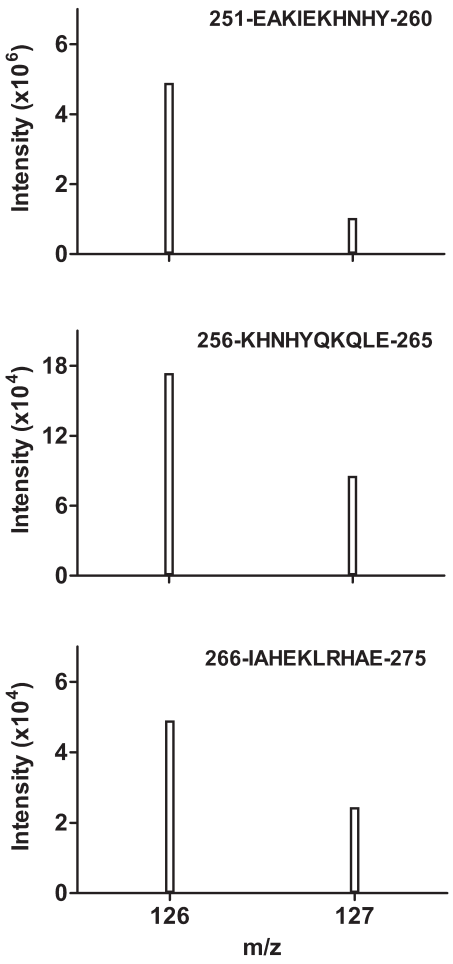


Figure 2: **Peptides including Lys256 and Lys270 exhibit reduced incorporation with TMT-127.** The RAP D3 domain was modified with TMT-126 in the absence of cluster II and with TMT-127 in the presence of cluster II. Proteins were mixed in a one-to-one molar ratio, cleaved into peptides, and analyzed by mass spectrometry. CID and HCD spectra are obtained of the TMT-modified peptide ions derived from the RAP D3 domain as described in materials and methods. Shown is part of HCD spectra that comprise the mass reported groups derived from TMT-126 and TMT-127 of the indicated peptides.

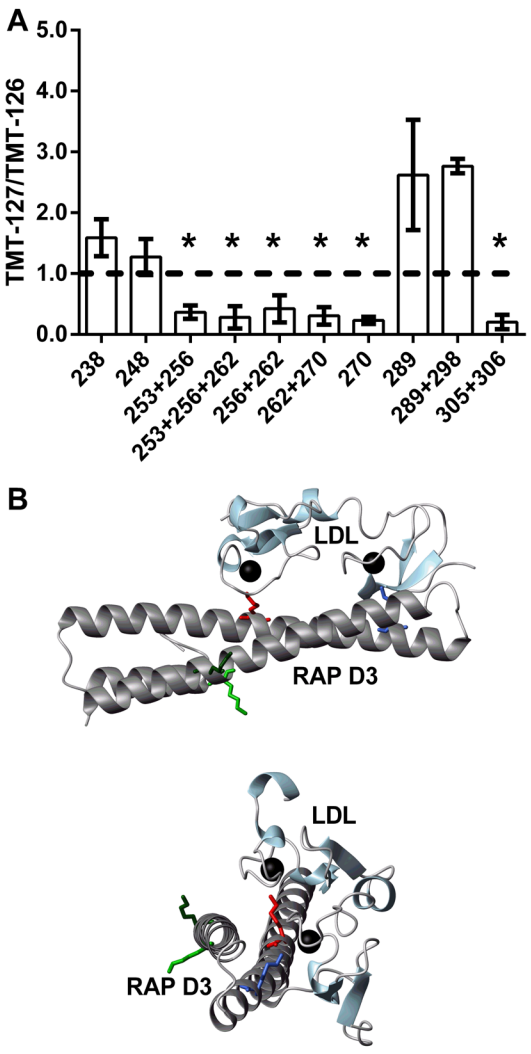


Figure 3: **Average TMT-127/TMT-126 ratio of the identified lysine containing peptides of the RAP D3 domain. (A)** The RAP D3 domain was modified with TMT-126 in the absence of cluster II and with TMT-127 in the presence of cluster II. Based on the concentration of the D3 domains, proteins were mixed in a one-to-one molar ratio, cleaved into peptides, and analyzed by mass spectrometry. The average TMT-127/TMT-126 ratio obtained from at least four independent experiments of peptides comprising the same lysine residues is displayed. Lysine residue numbers are indicated on the x-axis, those by which the average TMT-127/TMT-126 ratio is more than two standard deviations below 1 are indicated with an asterisk. **(B)** Two orientations of the crystal structure of the RAP D3 domain (in grey) in complex with two complement-type repeats of LDL receptor (in blue) (2FCW.pdb).[10] Indicated in red is Lys256 and in blue Lys270. Lys305 is shown in light green and Lys306 in dark green. The black spheres are calcium atoms that are critical for the structural integrity of the repeats.

residues 256 and/or 270. This suggests enhanced protection of these residues from modification by TMT-127 in the presence of cluster II implying that Lys256 and Lys270 contribute to the binding interaction. This is in full agreement with the crystal structure that shows that these residues contribute directly to the interaction with two consecutive complement-type repeats (Fig. 3B).[10] Intriguingly, we found that the peptides that include the lysine residues 305 and 306 exhibit a marked decrease in TMT-127/TMT-126 ratio as well (Fig. 3A). This suggests that either Lys305, Lys306 or both may contribute to the direct interaction with a third complement type repeat of LRP cluster II (Fig. 3B). Remarkably, the peptides including the lysine residues 238 or 289 show a marked increase in the incorporation of TMT-127 implying that

these residues have an increased surface exposure in the presence of cluster II. Taken together, next to Lys256 and Lys270 also the lysine residues at position 305 and/or 306 mediate RAP D3 domain binding to LRP cluster II.

Identification of critical LRP binding lysine residues within the D3 domain of full length RAP

We next evaluated whether we can identify the critical lysine residues of the D3 domain employing full length RAP of rat origin. Full length RAP of human and rat origin share 75% sequence identity and 87% sequence similarity (*Supplementary Figure S2*).[27] The RAP D3 domains share 81% sequence identity and 92% sequence similarity. All lysine residues of the human RAP D3 domain are conserved in rat RAP. Employing the footprinting approach, 94 peptides of full length RAP were identified covering 89% of the sequence of the protein. The TMT-modified peptides included 34 out of 37 lysine residues of RAP (*Supplementary Table S2*). The obtained average TMT-127/TMT-126 ratio of the lysine containing peptides derived from the D3 domain is shown in Figure 4C. The enhanced modification with TMT-127 of the lysine residues 238 and 289, which was observed for the isolated D3 domain in complex with cluster II (Fig. 3A), is not found employing the full length RAP-cluster II complex. However, the result does again reveal that Lys256 and Lys270 exhibit a reduced incorporation of TMT-127 (Fig. 4C). In addition, the peptide including Lys305 and Lys306 also shows a decreased TMT-127/TMT-126 ratio (Fig. 4C). This finding demonstrates that the residues of the D3 domain that contribute to LRP binding can be identified by our approach in both the isolated human D3 domain as well as in full length rat RAP.

Identification of novel LRP binding sites within RAP

Lys60 of the human RAP D1 domain (Fig. 4A, right panel) has been implicated to interact with a complement type repeat of cluster II.[13] No information is available about the lysine residues of the D2 domain that contribute to cluster II binding. In complete agreement with the NMR study, Figure 4A reveals a reduced incorporation of TMT-127 for the peptides including Lys60 of full length RAP from the rat (Fig. 4A). A small decrease in the TMT-127/TMT-126 ratio was also observed for the peptide including the lysine residues 63, 73 and 76. According to the NMR structure of RAP, these residues are, however, in close proximity of Lys60.[28] It seem therefore unlikely that these residues interact with a second complement-type repeat of cluster II. For the D2 domain, only the peptides including Lys191 showed a decreased TMT-127/TMT-126 ratio (Fig. 4B). Our data strongly suggests a previously unidentified role of Lys191 for cluster II binding.

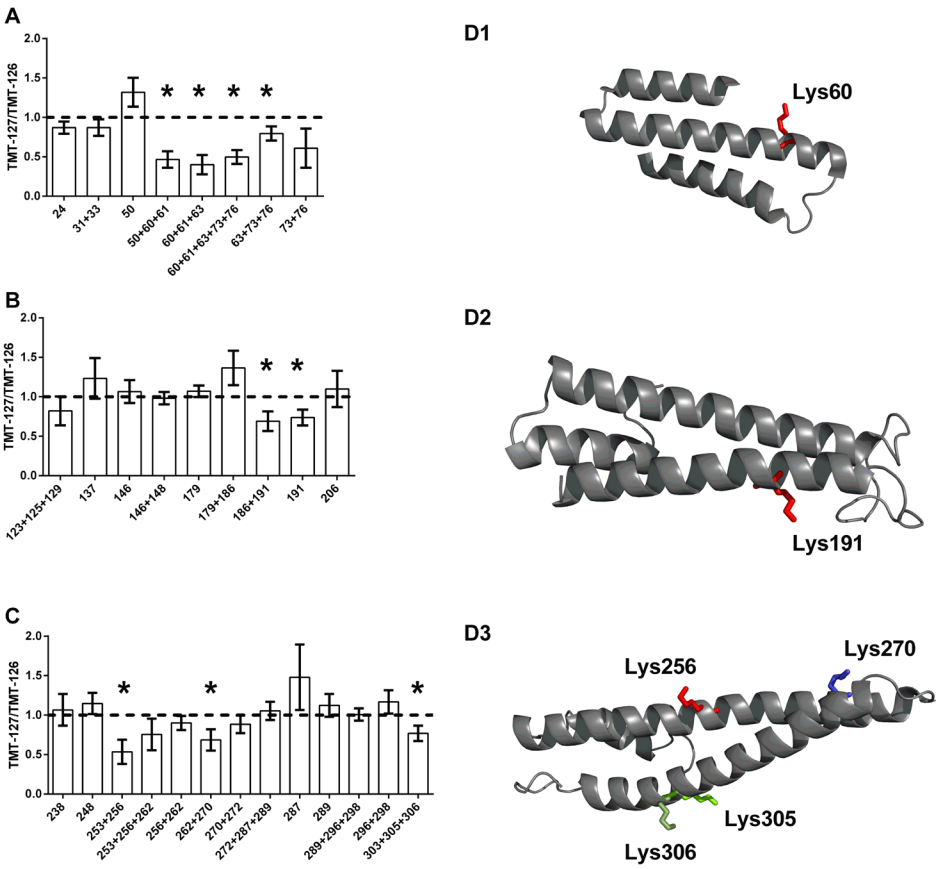


Figure 4: **Average TMT-127/TMT-126 ratio of the identified lysine containing peptides of full length RAP.** Full length RAP was modified with TMT-126 is the absence of cluster II and with TMT-127 in the presence of cluster II. Proteins were mixed in a one-to-one molar ratio, cleaved into peptides and analyzed by mass spectrometry. The average TMT-127/TMT-126 ratio obtained from at least four independent experiments of peptides comprising the same lysine residues is displayed. Lysine residue numbers are indicated on the x-axis. The top panel shows the lysine residues from the D1 domain of RAP, the middle panel the lysine residues from the D2 domain and the bottom panel the lysine residues from the D3 domain. Shown on the right are the NMR structures of the individual domains of RAP (2P03.pdb).[28] In panel A, Lys60 is displayed in red. In panel B, Lys191 is displayed in red. In panel C, Lys256 is displayed in red, Lys270 in blue, Lys305 and Lys306 in light and dark green, respectively. Lysine residues of which the average TMT-127/TMT-126 ratio is more than two standard deviations below 1 are indicated with an asterisk.

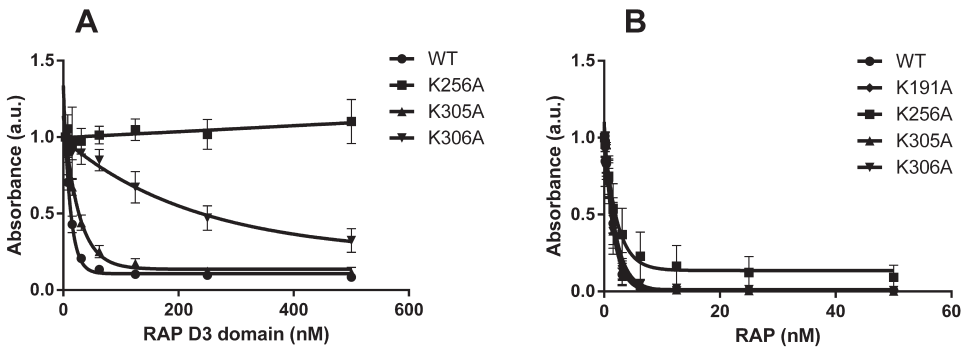


Figure 5: **The role of the identified lysine residues for the interaction with cluster II. (A)** Increasing concentrations of the indicated RAP D3 domain variants were incubated with 2.5 nM cluster II. The proteins were added to immobilized RAP D3 domain. Residual cluster II binding to immobilized D3 domain was assessed employing peroxidase labelled monoclonal antibody CLB-CAg69.[24,25] **(B)** Increasing concentrations of the indicated full length RAP variants were incubated with 0.5 nM cluster II. The proteins were added to immobilized full length RAP D3. Residual cluster II binding to immobilized D3 domain was assessed employing peroxidase labelled monoclonal antibody CLB-CAg69.

Lys306 contributes to the binding of the RAP D3 domain to cluster II

WT RAP and Lys191Ala, Lys256Ala, Lys305Ala, and Lys306Ala variants were employed in a ligand-binding competition assay to evaluate the role of the identified residues for RAP-cluster II complex formation. To this end, WT RAP was immobilized to a microtiter plate and incubated with cluster II to assess the concentration at which half-maximum binding is reached (data not shown). This cluster II concentration (i.e. 0.5 nM) was then incubated with immobilized WT RAP in the presence of increasing concentrations of RAP and the variants thereof (Fig. 5B). Results showed that all variants but one were equally effective in competing with immobilized RAP for binding cluster II. Only the Lys256Ala variant revealed a small decrease in the competition efficiency. Apparently, mutagenesis of a single contact site between RAP and cluster II has only a small effect on the binding interaction. We next assessed cluster II binding of the RAP D3 domain variants Lys256Ala, Lys305Ala, and Lys306Ala in a ligand-binding competition assay. Half-maximum binding of cluster II to the immobilized D3 domain was reached at a concentration of 2.5 nM. The competition data revealed that a concentration of 4 nM of WT D3 domain was required to obtain 50% residual binding of cluster II to immobilized RAP D3 domain (Fig. 5A). The D3 domain Lys256Ala variant was completely unable to inhibit the binding of cluster II to immobilized WT D3 domain. This agrees with the observation that Lys256 is critical for the interaction between WT D3 domain and a complement-type repeat. For Lys305Ala and Lys306Ala a

concentration of respectively 15 and 303 nM was required to reach a residual binding of 50%. This implies that mainly Lys306 contributes to LRP cluster II binding.

DISCUSSION

The crystal structure of the complex between RAP D3 domain and two-complement type repeats has increased our understanding about the mechanism of interaction between RAP and the LDL receptor family members. The negatively charged residues of the two complement-type repeats form an acidic necklace around the positively charged side chains of the lysine residues 256 and 270.[10] The present study demonstrates that TMTs can be employed to successfully identify these lysine residues (Fig. 2-4). Application of the approach also confirms the earlier established role of Lys60 for complex formation with a complement-type repeat. Our results together strongly suggest that the lysine residues Lys191 and Lys306 interact with the negatively charged acidic necklace of a complement-type repeat as well (Fig. 3-5). We therefore propose that Lys60, Lys191, Lys256, Lys270 and Lys306 are the critical lysine residues that directly contribute to complex formation between RAP and cluster II.

Chemical footprinting studies can be severely hampered by dissociation and re-association kinetics of a protein complex in solution. Critical residues can be modified upon dissociation of the complex, which may even lead to a conformational change of one of the binding partners. This will shift the binding equilibrium towards the unbound state of the proteins, which will further enhance TMT modification of the critical residues. It is therefore not surprising that the lysine residues that directly contribute to complex formation are also modified with TMT-127 in spite of the presence of cluster II (Fig. 2-4). It can further not be excluded that there is a dynamic equilibrium between distinct RAP-cluster II complexes in which one or more critical lysine residues are not occupied by a complement-type repeat. This may then explain why the difference in TMT ratios for full length RAP are less pronounced than those obtained for the much smaller isolated D3 domain. In spite of these notions, the difference in chemical modification is sufficiently large to successfully identify the critical lysine residues.

In an elegant study by Dolmer *et al*.[29], the interaction between two complement-type repeats and Lys253, 256, 289 and 270 of the D3 domain was evaluated by replacing all other lysine residues for alanines. They showed that Lys253 and Lys289 support the interaction of the critical lysine pair 256 and 270 with the two complement-type repeats. In our study, the intact LRP cluster II of 9 complement-type repeats in interaction with the D3 domain revealed that Lys289 is more accessible for

modification with TMT-127 in the protein complex. This implies that Lys289 is more exposed to the protein surface after complex formation. This suggests that Lys289 does not contribute to the direct interaction in presence of multiple complement-type repeats. The increased TMT-127/TMT-126 ratio at Lys289 is not observed upon cluster II binding to full length RAP from the rat (Fig. 4). This may be related to a difference in local structure of the D3 domain from rat and human origin. Alternatively, it cannot be excluded that Lys289 contributes to intra-domain interactions in full length RAP. This interaction may remain unaltered upon binding cluster II. If so, this will then also not lead to an increased solvent exposure after cluster II binding.

The crystal structure as well as the NMR structure of RAP show that Lys306 is at the same side of the D3 domain as Lys256 and Lys270 (Fig. 3B, Fig. 4C).[10,28] The role of Lys306 for cluster II binding opens, therefore, the possibility that three consecutive complement-type repeats of cluster II interact with the D3 domain of RAP. This is compatible with the observation that three complement-type repeats exhibit a higher binding affinity for the D3 domain than two complement type repeats.[20] For the D1 domain of RAP, Lys60 has been identified to contribute to complex formation in the present and previous studies (Fig. 4A).[13] This suggests that a single complement-type repeat interacts with the D1 domain via the acidic necklace binding mechanism. The positively charged residues Lys63, 73, and 76 are in close proximity to Lys60 and may therefore support this interaction. This may explain the small reduction in TMT ratio of peptides that include Lys63, 73, and 76. The D2 domain also seems to interact with only a single complement-type repeat via the acid necklace binding model as only Lys191 showed a reduced TMT ratio. These findings are compatible with the previous observations that the isolated D1 and D2 domains are less effective in binding cluster II fragments.[13,18]

The combined binding sites in RAP for the complement-type repeats mediate the particularly effective interaction with LRP. Our results suggest that even the absence of the critical residue Lys256 in full length RAP can be compensated by the other critical lysine residues (Fig. 5). The notion that these binding sites are distributed over multiple domains of RAP provides insight into the general mechanism by which LRP binds its ligand. The four clusters of complement type repeats in LRP may interact with a multitude of lysine residues that are distributed over a large area of the ligand. Pinpointing a single binding site for LRP on a ligand may therefore not be possible. This provides, for instance, an explanation why multiple LRP binding regions have been identified for coagulation factor VIII.[23,30]

The employed chemical footprinting approach can be applied to any protein complex assembly that involves critical lysine residues. The interaction of CUB domains with protein binding partners has, for instance, been suggested to involve lysine residues.[31] The identity of these lysine residues can now be reliably identified

with the described approach. Taken together, we have developed a powerful approach to identify critical lysine residues for effective complex formation between proteins.

Author contribution

Esther Bloem and Eduard Ebberink performed the research and wrote the paper. Maartje van den Biggelaar contributed to the experimental design of the study and assisted in the purification of the required proteins. Carmen van der Zwaan provided technical assistance. Koen Mertens provided expert advice for the overall study and contributed to editing of the paper. Alexander Meijer designed the research, provided guidance in the analysis and interpretation of the results, and contributed to writing, drafting, and editing of the paper.

Abbreviations

TMT, Tandem Mass Tags; RAP, Receptor-Associated-Protein; LRP, Low-density lipoprotein receptor-Related Protein; LDL, Low-Density Lipoprotein; LDLR, Low-Density Lipoprotein Receptor; CID, Collision Induced Dissociation; HCD, Higher energy Collision induced Dissociation; WT, wild-type.

REFERENCES

1. Berggard T, Linse S, James P. Methods for the detection and analysis of protein-protein interactions. *Proteomics.* 2007 **7** 2833-2842
2. Pacholarz, KJ, Garlish RA, Taylor RJ, Barran PE. Mass spectrometry based tools to investigate protein-ligand interactions for drug discovery. *Chem. Soc. Rev.* 2012 **41** 4335-4355
3. Guan J-Q, Chance MR. Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry. *Trends Biochem. Sci*. 2005 **30** 583-592
4. McKee CJ, Kessl JJ, Norris JO, Shkriabai N, Kvaratskhelia M. Mass spectrometry-based footprinting of protein-protein interactions. *Methods* 2009 **47** 304-307
5. Ori A, Free P, Courty J, Wilkinson MC, Fernig DG. Identification of heparin-binding sites in proteins by selective labeling. *Mol. Cell. Proteomics.* 2009 **8** 2256-2265
6. Bloem E, Meems H, van den Biggelaar M, van der Zwaan C, Mertens K, Meijer AB. Mass spectrometry-assisted study reveals that lysine residues 1967 and 1968 have opposite contribution to stability of activated factor VIII. *J. Biol. Chem.* 2012 **287** 5775-5783
7. Castro-Nunez L, Bloem E, Boon-Spijker MG, van der Zwaan C, van den Biggelaar M, Mertens K, Meijer AB. Distinct roles of Ser-764 and Lys-773 at the N terminus of von Willebrand factor in complex assembly with coagulation factor VIII. *J. Biol. Chem.* 2013 **288** 393-400
8. Zhou Y, Vachet RW. Covalent labeling with isotopically encoded reagents for faster structural analysis of proteins by mass spectrometry. *Anal. Chem*. 2013 **85** 9664-9670
9. Dayon L, Pasquarello C, Hoogland C, Sanchez JC, Scherl A. Combining low- and high-energy tandem mass spectra for optimized peptide quantification with isobaric tags. *J. Proteomics.* 2010 **73** 769-777
10. Fisher C, Beglova N, Blacklow SC. Structure of an LDLR-RAP Complex Reveals a General Mode for Ligand Recognition by Lipoprotein Receptors. *Mol. Cell* 2006 **22** 277-283
11. Migliorini MM, Behre EH, Brew S, Ingham KC, Strickland DK. Allosteric modulation of ligand binding to low density lipoprotein receptor-related protein by the receptor-associated protein requires critical lysine residues within its carboxyl-terminal domain. *J. Biol. Chem.* 2003 **278** 17986-17992
12. Van den Biggelaar M, Sellink E, Klein Gebbinck JWTM, Mertens K, Meijer AB. A single lysine of the two-lysine recognition motif of the D3 domain of receptor-associated protein is sufficient to mediate endocytosis by low-density lipoprotein receptor-related protein. *Int. J. Biochem. Cell Biol.* 2011 **43** 431-440

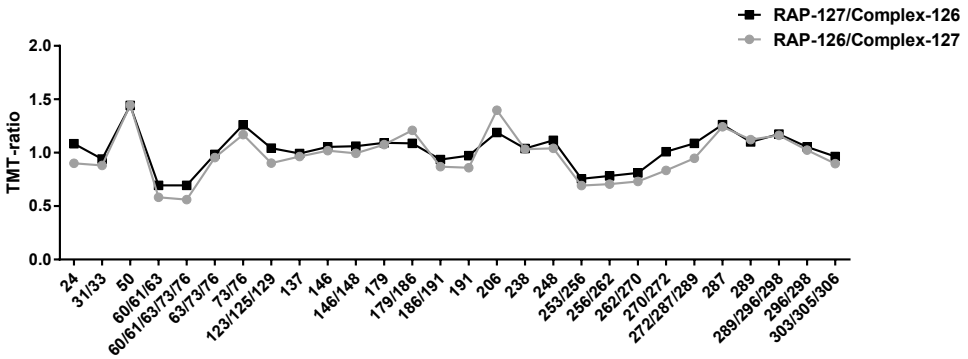
13. Jensen GA, Andersen OM, Bonvin AM, Bjerrum-Bohr I, Etzerodt M, Thøgersen HC, O'Shea C, Poulsen FM, Kragelund BB. Binding site structure of one LRP-RAP complex: implications for a common ligand-receptor binding motif. *J. Mol. Biol.* 2006 **362** 700-716
14. Herz J, Strickland DK. LRP: a multifunctional scavenger and signaling receptor. *J. Clin. Invest.* 2001 **108** 779-784
15. Willnow TE. Receptor-associated protein (RAP): a specialized chaperone for endocytic receptors. *Biol. Chem.* 1998 **379** 1025-1031
16. Herz J, Goldstein JL, Strickland DK, Ho YK, Brown MS. 39-kDa protein modulates binding of ligands to low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor. *J. Biol. Chem.* 1991 **266** 21232-21238
17. Herz J. The switch on the RAPper's necklace. *Mol. Cell* 2006 **23** 451-455
18. Andersen OM, Schwarz FP, Eisenstein E, Jacobsen C, Moestrup SK, Etzerodt M, Thøgersen HC. Dominant thermodynamic role of the third independent receptor binding site in the receptor-associated protein RAP. *Biochemistry* 2001 **40** 15408-15417
19. Estrada K, Fisher C, Blacklow SC. Unfolding of the RAP-D3 helical bundle facilitates dissociation of RAP-receptor complexes. *Biochemistry* 2008 **47** 1532-1539
20. Jensen JK, Dolmer K, Gettins PGW. Receptor-associated protein (RAP) has two high-affinity binding sites for the low-density lipoprotein receptor-related protein (LRP): consequences for the chaperone functions of RAP. *Biochem. J.* 2009 **421** 273-282
21. Lillis AP, Van Duyn LB, Murphy-Ullrich JE, Strickland DK. LDL receptor-related protein 1: unique tissue-specific functions revealed by selective gene knockout studies. *Physiol. Rev.* 2008 **88** 887-918
22. Zaiou M, Arnold KS, Newhouse YM, Innerarity TL, Weisgraber KH, Segall ML, Phillips MC, Lund-Katz S. Apolipoprotein E-low density lipoprotein receptor interaction. Influences of basic residue and amphipathic alpha-helix organization in the ligand. *J. Lipid Res.* 2000 **41** 1087-1095
23. Bovenschen N, Boertjes RC, Van Stempvoort G, Voorberg J, Lenting PJ, Meijer AB, Mertens K. Low density lipoprotein receptor-related protein and factor IXa share structural requirements for binding to the A3 domain of coagulation factor VIII. *J. Biol. Chem.* 2003 **278** 9370-9377
24. Stel HV, Sakariassen KS, Scholte BJ, Veerman EC, Van der Kwast TH, de Groot PG, Sixma JJ, van Mourik JA. Characterization of 25 monoclonal antibodies to factor VIII-von Willebrand factor: relationship between ristocetin-induced platelet aggregation and platelet adherence to subendothelium. *Blood* 1984 **63** 1408-1415
25. Meijer AB, Rohlena J, van der Zwaan C, van Zonneveld A, Boertjes RC, Lenting PJ,

Mertens K. Functional duplication of ligand-binding domains within low-density lipoprotein receptor-related protein for interaction with receptor associated protein, α2-macroglobulin, factor IXa and factor VIII. *Biochimica et Biophysica Acta* 2007 **1774** 714-722

26. Van Haren, SD, Herczenik E, ten Brinkem A, Mertens K, Voorberg J, Meijer AB. HLA-DR-presented peptide repertoires derived from human monocyte-derived dendritic cells pulsed with blood coagulation factor VIII. *Mol. Cell. Proteomics* 2011 **10** M110.002246
27. UniProt Consortium. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic. Acids Res.* 2012 **40** D71-5
28. Lee D, Walsh JD, Migliorini M, Yu P, Cai T, Schwieters CD, Krueger S, Strickland DK, Wang YX. The structure of receptor-associated protein (RAP). *Protein Sci.* 2007 **16** 1628-1640
29. Dolmer K, Campos A, Gettins PGW. Quantitative dissection of the binding contribution of ligand lysines of the receptor-associated protein (RAP) to the low density lipoprotein receptor-related protein (LRP1). *J. Biol. Chem.* 2013 **288** 24081-24090
30. Henriët M, van den Biggelaar M, Rondaij M, van der Zwaan C, Mertens K, Meijer AB. C1 domain residues Lys 2092 and Phe 2093 are of major importance for the endocytic uptake of coagulation factor VIII. *Int. J. Biochem. Cell Biol.* 2011 **43** 1114-1121
31. Andersen CB, Madsen M, Storm T, Moestrup SK, Andersen GR. Structural basis for receptor recognition of vitamin-B(12)-intrinsic factor complexes. *Nature* 2010 **464** 445-448

SUPPLEMENTARY SECTION

SUPPLEMENTARY FIGURE S1



Supplementary Figure 1: **Reversing the TMT labels in the footprinting strategy does not affect the outcome of the study.** RAP was modified with TMT-127 in the absence of cluster II and with TMT-126 in the presence of cluster II. The TMT-127/TMT-126 ratios of the lysine containing peptides are shown in the figure. The same approach was repeated but then RAP was modified with TMT-126 and the complex with TMT-127. The figure shows the TMT-126/TMT-127 ratios for the lysine containing peptides for this setup in the figure as well. Numbers of the involved lysine residues are indicated on the x-axis.

SUPPLEMENTARY FIGURE S2

Supplementary Figure 2: **Sequence alignment of full length RAP of human and rat origin.** The primary sequence of rat and human RAP were retrieved and compared employing the Uniprot online tools (ref 27). Indicated in red are the critical lysine residues that were identified.

SUPPLEMENTARY TABLE S1

Sequence	Activation Type	Xcorr	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]
215-GAEFEEPRVI-224	CID	2.7	2	573.799	1146.590	9.58	31.64
215-GAEFEEPRVIDLW-227	CID	3.39	2	780.893	1560.780	6.55	47.2
219-EEPRVIDLW-277	HCD	2.29	2	578.800	1156.593	-5.9	43.35
226-LWDLAQSANLTDkE-239	CID	4.11	2	914.987	1828.968	8.44	37.49
228-DLAQSANLTDkEL-240	CID	3.53	2	821.937	1642.868	-3.17	35.02
240-LEAFREE-246	CID	1.62	2	447.223	893.438	1.59	28.46
241-EAFREELkHF-250	CID	2.31	3	510.943	1530.816	0.91	32.71
242-AFREELkHFE-251	CID	2.52	4	383.460	1530.817	1.57	32.07
244-REELkHF-250	CID	1.86	2	592.337	1183.667	0.16	27.13
247-LkHFE-251	CID	1.82	3	300.178	898.520	-3.12	22.78
247-LkHFEAkIE-255	CID	2.33	3	522.315	1564.931	-3.75	26.90
251-EAkIEkHNHY-260	HCD	2.34	4	430.496	1718.961	6.5	21.41
252-AkIEkHNHYQkQLE-265	HCD	4.07	4	611.106	2441.403	0.32	26.59
256-kHNHYQkQLE-265	CID	3.35	3	592.339	1775.001	7.77	20.88
261-QkQLEIAHEkL-271	HCD	3.97	3	596.362	1787.071	0.58	30.89
265-EIAHEkL-271	CID	1.29	3	355.542	1064.613	-5.16	27.3
266-IAHEkLRHAE-275	CID	2.53	3	476.939	1428.804	-7.89	21.74
272-RHAESVGDGERVSRSREkHALL-293	CID	2.63	5	543.703	2714.486	8.2	22.08
289-kHALLE-294	CID	1.58	2	468.292	935.578	2.36	59.97
289-kHALLEGRTkE-299	HCD	3.07	3	578.015	1732.030	-5.39	22.97
300-LGYTVkkHLQDLSGRISRARHNEL-323	CID	2.38	6	541.150	3241.862	1.85	63.94
303-TVkkHL-308	CID	1.54	3	392.595	1175.772	-5.78	24.98
303-TVkkHLQDL-311	HCD	2.31	4	383.742	1531.948	0.01	28.67
310-DLSGRISRARHNEL-323	HCD	1.8	2	812.445	1623.883	9.41	24.12
311-LSGRISRARHNEL-323	CID	1.35	4	377.966	1508.843	1.36	22.71
312-SGRISRARHNEL-323	CID	2.89	3	465.920	1395.746	-7.19	20.27

Supplementary Table S1: **Identified peptides derived from the D3 domain of RAP.** Peptides of the D3 domain of RAP with the best Xcorr are displayed. Shown is amino acid numbering, Xcorr, charge state, mass/charge (m/z), protonated mass (MH^+), mass deviation from the theoretical mass (ΔM) and retention time (RT).

SUPPLEMENTARY TABLE S2

Sequence	Activation Type	XCorr	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]
16-SGEEFRME-23	CID	2.54	2	492.709	984.411	1.35	29.76
21-RMEkL-25	HCD	1.35	3	301.184	901.538	1.35	25.91
21-RmEkLNQLW-29	HCD	1.23	3	486.938	1458.798	0.95	35.05
21-RMEkLNQLW-29	HCD	2.08	3	481.605	1442.799	-1.76	36.42
24-kLNQLWE-30	CID	1.50	2	578.332	1155.657	-3.07	36.71
26-NQLWEkAkRL-35	HCD	1.64	4	434.767	1736.047	-1.50	34.24
30-EkAkRL-35	HCD	0.84	3	398.931	1194.779	-4.72	24.61
31-kAkRLHLSPVRLAE-44	HCD	3.95	5	414.467	2068.308	2.52	35.75
36-HLSPVRL-42	HCD	1.70	2	411.254	821.500	0.92	26.18
36-HLSPVRLAEL-45	CID	1.83	2	567.835	1134.664	0.42	34.12
38-SPVRLAEL-45	CID	1.50	2	442.764	884.520	0.24	33.89
38-SPVRLAELHSDL-49	CID	3.53	3	446.246	1336.724	1.15	33.46
42-LAELHSDLkIQER-54	CID	3.41	3	593.005	1776.999	-3.18	29.83
45-LHSDLkIQE-53	CID	2.83	2	654.374	1307.742	1.30	29.18
46-HSDLkIQERDELNW-59	CID	3.11	3	670.017	2008.037	1.84	56.32
48-DLkIQER-54	CID	2.98	3	376.225	1126.661	-4.86	27.87
49-LkIQE-53	CID	0.92	2	428.271	855.535	-3.58	25.6
50-kIQERDEL-57	CID	2.45	3	419.241	1255.710	0.94	25.97
50-kIQERDELNW-59	CID	2.80	2	778.420	1555.832	0.92	32.78
50-kIQERDELNWkkL-62	CID	2.74	5	475.888	2375.409	-3.35	29.28
55-DELNWkkLkVEGL-67	HCD	2.47	4	562.591	2247.343	-1.81	39.56
57-LNWkkLkVE-65	HCD	2.29	4	459.048	1833.169	-1.54	34.47
60-kkLkVEGLDGDGEkEAkL-77	HCD	4.24	6	514.653	3082.879	0.99	32.19
63-kVEGLDGDGEkEAkL-77	CID	4.46	4	566.578	2263.292	0.71	32.35
68-DGDGEkEAkL-77	CID	1.95	3	504.609	1511.813	-6.39	30.28
68-DGDGEkEAkLVHNLNVILARYGL-90	CID	1.72	5	595.738	2974.659	-1.46	49.99
78-VHNLNVIL-85	HCD	1.46	2	461.280	921.553	1.22	35.99
78-VHNLNVILARY-88	HCD	2.93	3	437.921	1311.747	-4.50	35.45
82-NVILARY-88	CID	1.34	2	424.753	848.498	-1.01	31.71
91-DGRkDTQTVHSNALNEDTQ-109	HCD	1.85	4	589.288	2354.128	-5.51	25.43
95-DTQTVHSNALNE-106	CID	2.27	2	664.807	1328.606	-1.31	24.28
95-DTQTVHSNALNEDTQ-109	CID	3.87	2	836.870	1672.732	-5.51	25.92
119-kLWHkAkTSGkFSSEE-134	HCD	2.47	5	553.524	2763.593	-2.30	27.92
121-WHkAkTSGkF-130	CID	2.30	4	467.035	1865.116	0.68	27.08
122-HkAkTSGkF-130	HCD	3.15	4	420.515	1679.038	1.54	24.55
130-FSSEELDkLWR-140	CID	2.18	3	545.626	1634.864	1.60	39.53
131-SSEELDkL-138	CID	1.93	2	573.307	1145.606	-6.10	33.88
131-SSEELDkLW-139	CID	1.63	2	666.350	1331.693	0.82	40.57
135-LDkLWRE-141	CID	2.49	3	395.564	1184.678	-7.58	33.35
140-REFLHY-145	CID	1.57	2	432.722	864.436	-0.58	29.15
142-FLHYkE-147	CID	2.28	3	354.532	1061.581	-4.82	29.32
143-LHYkEkIHEY-152	CID	2.55	3	604.009	1810.013	-2.39	28.2
146-kEkIHEY-152	CID	1.97	4	349.957	1396.808	-2.02	25.04
152-YNVLLDTLSRAEE-164	HCD	3.66	2	761.890	1522.773	-1.44	43.34
157-DTLSRAEEGYENLLSPS-173	CID	3.11	2	940.946	1880.884	-1.84	39.98
157-DTLSRAEEGYENLLSPSDMTHIkS-180	HCD	5.40	4	730.364	2918.434	1.88	40.29
167-ENLLSPSDMTHIkSDTL-183	CID	2.05	3	709.369	2126.091	1.61	39.39
168-NLLSPSDMTHIkSDTLASkHSE-189	HCD	5.67	5	573.104	2861.493	-2.14	35.57
174-DMTHIkS-180	HCD	1.22	3	352.857	1056.557	-1.36	26.59
181-DTLASkHSELk-191	CID	3.47	3	560.327	1678.966	0.75	27.85
181-DTLASkHSELkDRLRSINQGL-201	CID	1.24	5	567.127	2831.604	0.89	37.01
184-ASkHSEL-190	HCD	0.89	3	332.857	996.558	2.50	22.59
184-ASkHSELkDRL-194	HCD	2.28	4	434.261	1734.021	1.40	26.7
191-kDRLRSINQGLDRL-203	HCD	3.79	4	478.035	1909.119	0.77	31.35
192-DRLRSINQGL-201	HCD	2.03	2	586.329	1171.651	-2.53	29.07
195-RSINQGLDRL-204	CID	2.13	2	586.331	1171.654	-0.02	28.78
202-DRLRkVSHQGY-212	CID	2.39	4	396.726	1583.883	-0.91	22.9
202-DRLRkVSHQGYGPATEF-218	HCD	3.11	4	547.292	2186.148	-3.02	28.57
202-DRLRkVSHQGYGPATEFEEPRVI-223	HCD	5.42	4	728.140	2909.540	-2.18	31.25
203-RLRkVSHQGYGPATE-217	HCD	2.29	4	481.770	1924.057	-1.38	22.99
205-RkVSHQGY-212	HCD	1.52	3	400.563	1199.675	1.98	19.49
205-RkVSHQGYGPATEF-218	HCD	2.24	3	601.319	1801.941	-0.72	27.55

Sequence	Activation Type	XCorr	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]
213-GPATEFEEPRVIDLW-227	CID	4.79	2	879.940	1758.873	1.79	45.16
218-FEEPRVID-225	CID	1.79	2	502.755	1004.503	-2.23	31.53
219-EEPRVIDLW-227	CID	2.11	2	578.804	1156.601	1.07	42.18
223-VIDLWDLAQSANFTEkELESFR-244	CID	4.96	3	946.160	2836.464	1.57	53.32
226-LWDLAQSANFTEkE-239	CID	4.46	2	938.982	1876.957	2.57	41.47
228-DLAQSANF-235	CID	1.51	2	433.206	865.404	-1.46	33.87
228-DLAQSANFTEkEL-240	CID	3.00	2	845.942	1690.877	2.19	38.74
236-TEkELESF-243	CID	1.87	2	604.317	1207.626	-2.27	33.57
236-TEkELESFREEL-247	CID	3.75	3	578.969	1734.891	-4.55	32.02
243-SFREELkHFE-251	CID	2.00	4	387.457	1546.806	-2.17	29.29
244-REELkHF-250	CID	1.58	3	395.225	1183.660	-5.37	27.54
247-LkHFE-251	CID	1.77	3	300.177	898.518	-5.26	22.72
251-EAkIEkHNHY-260	HCD	2.33	3	573.655	1718.949	-0.38	23.75
251-EAkIEkHNHYQkQL-264	CID	1.37	5	489.085	2441.396	-2.37	25.84
252-AkIE-255	CID	1.29	2	343.218	685.429	-5.63	20.70
252-AkIEkHNHYQkQLE-265	HCD	2.99	6	407.739	2441.398	-1.81	25.05
256-kHNHYQkQLE-265	HCD	2.99	4	444.502	1774.985	-1.25	22.72
261-QkQLEISHQkL-271	CID	4.09	3	601.365	1802.081	0.12	29.58
265-EISHQkL-271	CID	1.38	2	540.318	1079.628	-0.52	25.41
266-ISHQkLkHVE-275	HCD	3.16	4	418.008	1669.009	0.99	23.72
272-kHVESIGDPEHISRNkEkY-290	HCD	4.55	6	491.111	2941.629	1.47	60.05
276-SIGDPEHISRNkE-288	HCD	2.87	4	483.767	1932.048	1.08	28.47
280-PEHISRNkE-288	CID	1.88	4	334.438	1334.729	2.69	20.54
288-EkYVLLEEkTk-298	CID	0.99	3	685.753	2055.245	-0.31	32.60
289-kYVLLE-294	CID	1.60	2	495.308	989.608	-3.38	31.36
289-kYVLLEE-295	CID	1.53	2	559.830	1118.652	-1.52	35.65
289-kYVLLEEkTkE-299	CID	3.12	3	685.754	2055.248	1.20	32.44
291-VLLEEkTkEL-300	HCD	2.26	3	551.344	1652.018	1.24	34.73
295-EkTkE-299	CID	1.86	3	362.220	1084.646	-5.50	19.79
303-kVkkHLQDL-311	HCD	1.77	4	446.795	1784.157	3.43	27.35
310-DLSSRVSRARHNEL-323	HCD	1.58	3	547.291	1639.860	-1.69	25.23
312-SSRVSRARHNEL-323	CID	1.75	3	471.256	1411.753	1.00	14.18

Supplementary Table S2: **Identified peptides derived from rat RAP.** Identified peptides of full length RAP with the best Xcorr are displayed. Shown is amino acid numbering, Xcorr, charge state, mass/charge (m/z), protonated mass (MH^+), mass deviation from the theoretical mass (ΔM) and retention time (RT). The D1 domain is defined from 1-112, the D2 domain from 113-218 and the D3 domain from 219-323. Peptides including lysine residues 5 and 13 were not recovered and TMT-data is missing from lysine residue 94.

3

Factor VIII/V C-domain swaps reveal discrete C-domain roles in factor VIII function and intracellular trafficking

Eduard H.T.M. Ebberink[1]*, Eveline A.M. Bouwens[1]*, Esther Bloem[1], Mariëtte G. Boon-Spijker[1], Maartje van den Biggelaar[1,2], Jan Voorberg[1,3], Alexander B. Meijer[1,2] and Koen Mertens[1,2]

Haematologica. 2017 102 (4):686-694

[1]Department of Plasma Proteins, Sanquin Research, 1066 CX Amsterdam, The Netherlands
[2]Pharmaceutical Sciences, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands
[3] Landsteiner Laboratory of AMC and Sanquin, University of Amsterdam, 1105 AZ Amsterdam, the Netherlands

**authors contributed equally*

ABSTRACT

Factor VIII C-domains are believed to contain specific function in cofactor activity and in interaction with von Willebrand factor. We have previously shown that factor VIII is co-targeted with von Willebrand factor to the Weibel-Palade bodies in blood outgrowth endothelial cells, even when factor VIII carries mutations in the light chain that are associated with defective von Willebrand factor binding. In this study, we addressed the contribution of individual factor VIII C-domains in intracellular targeting, von Willebrand factor binding and cofactor activity by factor VIII/V C-domain swapping.

Blood outgrowth endothelial cells were transduced with lentivirus encoding factor V, VIII or YFP-tagged C-domain chimeras, and examined by confocal microscopy. The same chimeras were produced in HEK293-cells for *in vitro* characterization and chemical footprinting by mass spectrometry.

In contrast to factor VIII, factor V did not target to Weibel-Palade bodies. The chimeras showed reduced Weibel-Palade body targeting, suggesting that this requires the factor VIII C1-C2 region. The factor VIII/V-C1 chimera did not bind von Willebrand Factor and had reduced affinity for factor IXa, whereas the factor VIII/V-C2 chimera showed a minor reduction in von Willebrand factor binding and normal interaction with factor IXa. This suggests that mainly the C1-domain carries factor VIII-specific features in assembly with von Willebrand factor and factor IXa. Footprinting analysis of the chimeras revealed increased exposure of lysine residues in the A1/C2- and C1/C2-domain interface, suggesting increased C2-domain mobility and disruption of the natural C-domain tandem pair orientation. Apparently, this does affect intracellular trafficking, but not extracellular function.

INTRODUCTION

Factor VIII (FVIII) serves as a co-factor for activated factor IX (FIXa) in the factor X (FX) activating complex. It consists of 2332 amino acids with a distinct domain structure A1-*a1*-A2-*a2*-B-*a3*-A3-C1-C2.[1] Intracellular processing of the B-domain yields a heterodimeric FVIII protein with a 90-220 kDa heavy chain (A1-*a1*-A2-*a2*) non-covalently associated with a 80 kDa light chain (*a3*-A3-C1-C2).[2] FVIII circulates in complex with the multimeric glycoprotein von Willebrand factor (VWF) that protects FVIII from premature clearance and proteolytic degradation. Complex assembly occurs over an extended surface on FVIII, spanning the entire light chain.[3-5] The sulfated tyrosine on position 1680 is essential for binding to VWF and mutation of this tyrosine results in impaired complex formation with VWF.[3]

Recently it has been shown that FVIII is expressed in endothelial cells.[6-8] Previous work showed that FVIII overexpressed in endothelial cells co-sorts with VWF to the secretory organelles designated Weibel-Palade bodies (WPBs).[9-11] The precise interaction mediating sorting to WPBs has not been clarified, although it has been generally assumed that VWF plays a key role as sorting chaperone. In contrast to this view, we have shown that FVIII sorting to WPBs does not require the high-affinity interaction via the sulfated tyrosine on position 1680 in the *a3*-domain.[11,12] Moreover, mutations in the FVIII C1- and C2-domains leading to impaired extracellular VWF/FVIII complex assembly show apparently normal expression of FVIII and storage in WPBs.[12]

Likewise to FVIII, coagulation factor V (FV) comprises two lipid-binding C-domains that form a similar side-by-side pair.[13-17] FV shares ~40% sequence homology with FVIII and has a similar domain structure (A1-A2-*a2*-B-A3-*a3*-C1-C2).[18] FV functions as a co-factor for FXa in the prothrombinase complex, demonstrating that FVIII and FV serve a similar co-factor function. Unlike FVIII however, FV neither circulates in complex with VWF, nor does it act as a cofactor for FIXa in the activation of FX. In the present study, we addressed the C-domains contribution to FVIII intracellular targeting and extracellular function by constructed FVIII chimeras carrying FV C1- or C2-domains and explored the functional and structural implications of these C-domain swaps.

METHODS

Factor VIII constructs

All constructs used in this study encoded B-domain-deleted FVIII (BDD-FVIII) variants in order to meet size restrictions in the lentiviral packaging system.[10] For the same reason FV was B-domain-deleted (BDD-FV) as well.[19] In BDD-FVIII-YFP, yellow fluorescent protein (YFP) replaced the B-domain, as described for its green fluorescent protein (GFP) equivalent elsewhere.[10,11] Construction of plasmids encoding the YFP-tagged BDD-FVIII/FV chimeras is described in the *Supplementary Section*. For functional studies, BDD-FVIII-YFP and the chimeras were constructed in the pcDNA3.1 vector for production in HEK293 cells. To simplify nomenclature, the term BDD-FVIII is replaced by FVIII throughout this paper. Consequently, the BDD-FVIII-YFP chimeras containing the FV-C1 or -C2 domain are referred to as FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2, respectively.

Immunofluorescence microscopy of lentiviral-transduced endothelial cells

Blood outgrowth endothelial cell (BOEC) isolation and subsequent transduction with lentivirus has been described previously.[10] A detailed description of the antibody staining of BDD-FV and BDD-FVIII can be found in the *Supplementary Section*. Z-stacks (0.4-µm intervals) were taken with confocal laser scanning microscopy using a Zeiss LSM510 equipped with Plan NeoFluar 63x/1.4 Oil objective (Carl Zeiss, Heidelberg, Germany). Images were processed with Zeiss LSM510 version 4.0 software and LSM image browser (Carl Zeiss, Heidelberg, Germany). Secretion of FVIII and FV was quantified by enzyme-linked immunosorbent assay (ELISA) as described previously, with the exception that the FV ELISA used the monoclonal anti-light chain antibody CLB-FV-4, and purified FV as a reference.[12,19]

Purified FVIII-YFP variants

FVIII-YFP/FV chimeras and FVIII-YFP were produced in stable cell lines (HEK293) and purified by immunoaffinity chromatography using a monoclonal antibody (VK34) followed by anion exchange chromatography (Q Sepharose FF, GE Healthcare, Uppsala, Sweden) as described in detail elsewhere.[20] Purified FVIII-YFP variants were homogeneous, and comprised a YFP-carrying heavy chain of approx. 110 kDa and a 80 kDa light chain (see *Supplementary Figure S1*). FVIII-YFP concentrations were determined by ELISA, employing the monoclonal anti-light chain antibody KM33 (anti-C1) or EL14 (anti-C2) for immobilization, and the anti-heavy chain antibody CLB-CAg-9 for detection.[21] Normal human plasma served as standard. FVIII activity was determined using the chromogenic assay (Chromogenix, Milan, Italy), and the

activity/antigen ratios were 1.0 for FVIII-YFP, 0.9 for FVIII-YFP/FV-C2, and 0.4 for FVIII-YFP/FV-C1. In all functional studies FVIII concentrations were based on antigen concentrations, assuming that 1 U/ml corresponds with 0.3 nM.

Characterization of FVIII-YFP variants

Interaction of purified FVIII-YFP variants to recombinant full-length VWF[11] was assessed by Surface Plasmon Resonance (SPR) analysis using a BIAcore 3000 biosensor (Biacore AB, Uppsala Sweden) as described previously.[21] Details on data analysis are provided in the *Supplementary Section*. Interaction of FVIII-YFP and the

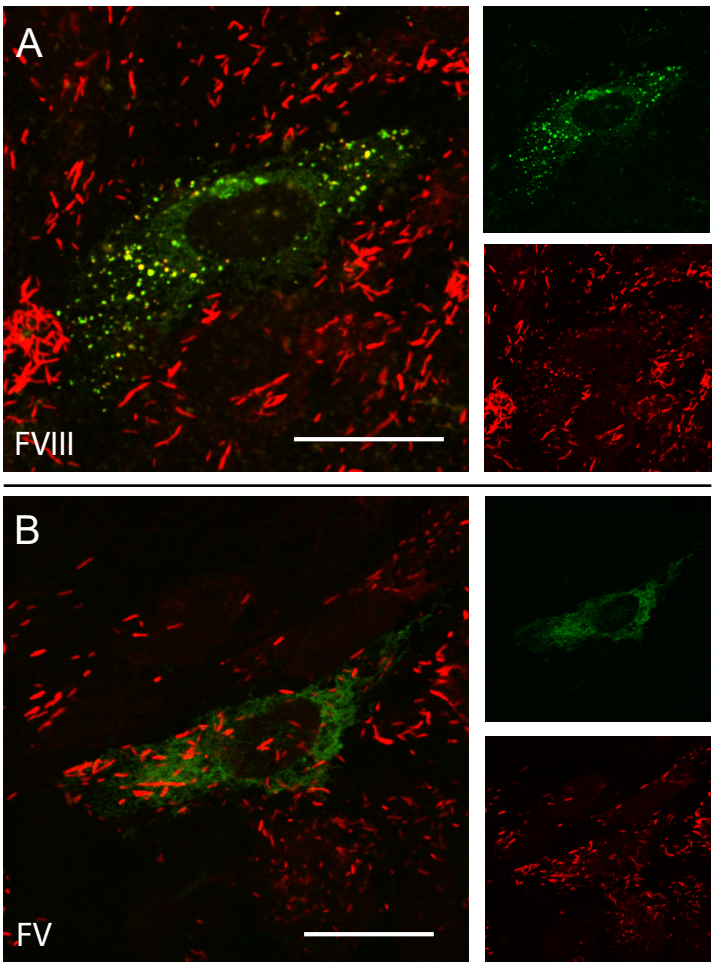


Figure 1: **Differential sorting of FVIII and FV in blood outgrowth endothelial cells.** Confocal images of BOECs expressing **(A)** B-domain-deleted FVIII or **(B)** B-domain-deleted FV. Both FVIII and FV are shown in green, and staining of VWF is shown in red. Yellow indicates co-localization of VWF with FVIII or FV. Images of the separate green and red channels are depicted on the right side. FVIII is solely visible co-stored with VWF in WPBs. While FVIII-containing WPBs have a round morphology, FVIII-negative WPBs remain elongated. FV is not visible in WPBs. The white scale bar represents 20 µm.

FVIII-YFP/FV chimeras with FIXa and phospholipids were inferred from FX activation studies as described in detail elsewhere.[21] Structural differences between FVIII-YFP, FV and the C-domain swapped chimeras were probed by chemical footprinting using lysine-reactive Tandem-Mass-Tags (TMTs) and mass spectrometry as described previously.[22] A full description of processing of labeled proteins into peptides and mass spectrometry analysis is given in the *Supplementary Section*.

RESULTS

Differential intracellular accumulation of FV compared to FVIII

Because FV is structurally highly homologous to FVIII, we compared FVIII and FV with respect to WPB trafficking in BOECs expressing BDD-FVIII or BDD-FV via lentiviral transduction. Transduced BOECs secreted much larger amounts of FV (140 pmol/10^6 cells/72 h) than FVIII (typically 1-5 pmol/10^6 cells/72 h). Confocal microscopy revealed that, as expected, FVIII retained within the endothelial cells completely co-localized with VWF in WPBs (Fig. 1A). Consistent with co-localization, FVIII-containing WPBs were round.[10,23,24] In contrast, FV did not traffic to WPBs (Fig. 1B). The FV-transduced cell contained WPBs which were negative for FV and had retained their typical elongated morphology similar to surrounding non-transduced cells. Instead, prominent background staining for FV could be detected in what might represent the endoplasmic reticulum (ER) (Fig. 1B). Apparently, trafficking of FV is different from that of FVIII.

C-domains of FVIII contribute to intracellular trafficking

To study the contribution of individual C-domains of the C-domain pair to FVIII sorting, we exchanged single FVIII C-domains for those of FV. Because of potential difficulties in staining of these variants with antibodies directed to the FVIII light chain, we expressed YFP-tagged FVIII (FVIII-YFP) in this experiment. To this end BOECs were transduced with lentivirus encoding for FVIII-YFP and YFP-tagged FVIII variants with swapped C1- or C2-domain (FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2). FVIII-YFP/FV chimera production as assessed by antigen levels in the conditioned medium ranged between 0.03 – 0.3 pmol/10^6 cells/72 h. The inherent variability in expression levels and transduction efficiency did not allow for quantification of the WPB-targeted fraction of the FVIII-YFP/FV chimeras. However, for transduced BOECs, trafficking could be studied by fluorescence microscopy. Confocal microscopy showed that YFP-tagged FVIII sorted exclusively to WPBs (Fig. 2A) and was retained in round WPBs, similar to untagged FVIII (Fig. 1A). This was to be expected because YFP-tags as such do not affect FVIII trafficking to WPBs.[12,24] Substitution of the C1-domain resulted in total loss

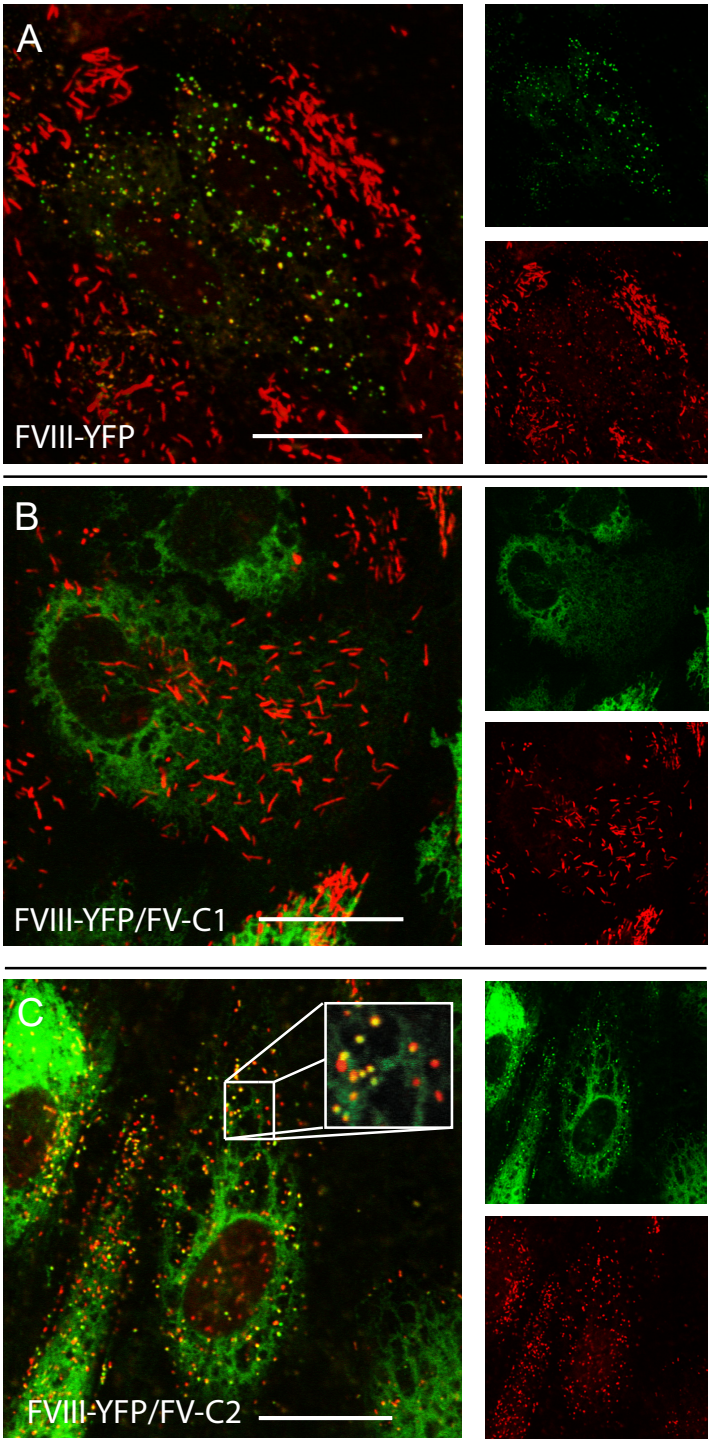


Figure 2: **Intracellular localization of FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2 chimeras in blood outgrowth endothelial cells. Confocal microscopy of BOECs expressing YFP-tagged B-domain-deleted FVIII variants.** Shown on the left side are merged signals of the YFP-tagged FVIII variants (green) and Alexa-633 stained VWF (red) (co-localization in yellow). On the right side image exports of the separate channels are displayed: YFP (green) top panel and Alexa-633 (red) lower panel. **(A)** YFP-tagged FVIII co-localizes with VWF in round WPBs, while surrounding non-transduced BOECs, which are negative for YFP, contain elongated WPBs. **(B)** FVIII-YFP/FV-C1 does not co-localize with VWF, and the WPBs remain elongated. **(C)** Part of FVIII-YFP/FV-C2 is located in round WPBs with VWF, but most of the FVIII-YFP/FV-C2 is dispersed throughout the cell. The white scale bar represents 20 µm.

of co-localization of FVIII-YFP/FV-C1 with VWF in WPBs (Fig. 2B). Like FV, FVIII-YFP/FV-C1 was shown to accumulate intracellularly, without any apparent punctuated YFP signal as seen in Figure 2A (see upper right side panels). The FVIII-YFP/FV-C2 variant displayed some residual sorting involving round WPBs (Fig. 2C). However, not all WPBs were positive for YFP (see inset, Fig. 2C), and a major part of the YFP signal was localized intracellularly, presumably in the ER. These results indicate that the C1-domain is the main driver of FVIII trafficking to WPBs, although the C2-domain contributes to this process as well.

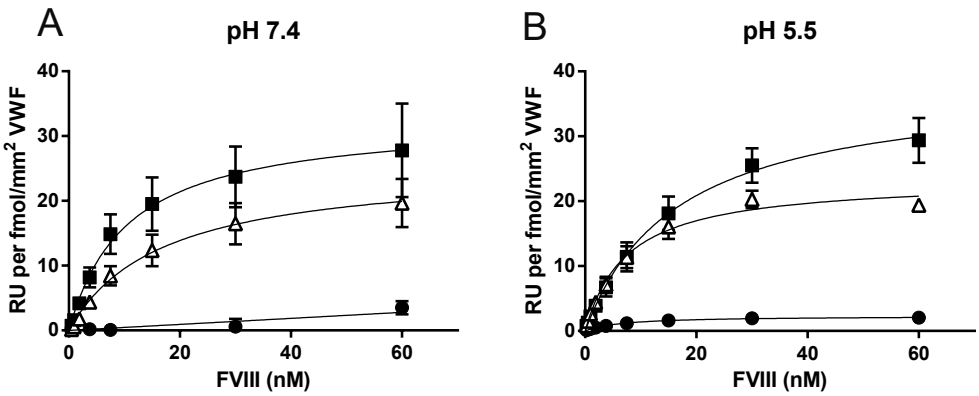


Figure 3: **Interaction of FVIII-YFP and FVIII-YFP/FV chimeras with VWF.** Surface Plasmon Resonance (SPR) analysis was performed using the BIAcore 3000 system as described.[21] FVIII-YFP (closed squares), FVIII-YFP/FV-C1 (closed circles) and FVIII-YFP/FV-C2 (open triangles) were passed over a CM5 chip coated with recombinant VWF (7, 24 and 37 fmol/mm^2) in a buffer containing 150 mM NaCl, 5 mM $CaCl_2$, 2.4% glycerol (v/v), 0.005% Tween 20 (v/v), and 20 mM HEPES (pH 7.4, panel A) or 20 mM MES (pH 5.5, panel B) for 240 seconds at 20 µL/min at 25 °C. The signal of a non-coated CM5 channel was subtracted to correct for differences in buffer composition. Response upon the onset of dissociation was taken to represent maximal binding and was plotted against the FVIII concentration. Values represent mean response units (RU) ± SD. Data were analyzed by a non-linear regression using a single hyperbola. This revealed apparent K_d values of approximately 8 and 16 nM for FVIII-YFP and FVIII-YFP/FV-C2, respectively. VWF binding of FVIII-YFP/FV-C1 was too low for quantitative analysis. Individual SPR sensorgrams are given in *Supplementary Figure S2*.

The FVIII C1-domain is critical for VWF-binding

Given that the FVIII-YFP/FV chimeras displayed reduced co-localization with VWF in BOECs, we studied the ability of the chimeras to interact with VWF employing Surface Plasmon Resonance (SPR) analysis. This enables a time-resolved monitoring of the association and dissociation between two interactive proteins, by measuring

mass increase and decrease due to the interaction of a soluble component (in this case, FVIII-YFP/FV chimeras) with an immobilized binding partner (in this case, VWF).[12,21] FVIII-YFP variants, in varying concentrations, were passed over a chip with immobilized VWF and analyzed for surface-bound mass change as expressed in Resonance Units (RU). Figure 3 shows the maximal binding response as a function of the FVIII concentration, which can be used to derive an estimate of the dissociation constant K_d.[21,25] The apparent K_d for FVIII-YFP was 8 nM, while FVIII-YFP/FV-C2 showed a slight reduction in VWF binding, with an apparent K_d of approximately 16 nM. In contrast, FVIII-YFP/FV-C1 displayed no appreciable interaction with VWF at all (Fig. 3A). Similar data were obtained pH 5.5, which should resemble the pH in the mature secretory compartment.[26] FVIII variants revealed complex binding kinetics, from which the affinity could not be directly inferred by standard curve fitting of the sensorgrams, due to a prominent heterogeneity in the dissociation phase at pH 5.5 (*Supplementary Figure S2*). Data obtained at pH 7.4 were less complex, and revealed an apparent K_d of 3 nM for FVIII-YFP and 8 nM for FVIII-YFP/FV-C2. Irrespective of the data analysis used, these experiments demonstrate that FVIII-YFP/FV-C2 and FVIII-YFP display similar interaction with VWF, whereas FVIII-YFP/FV-C1 shows almost no VWF binding. Apparently, replacement of the FVIII C2-domain with that of FV conserves VWF-binding. In contrast, the FVIII C1-domain is irreplaceable for interaction with VWF.

Substitution of FVIII C1-domain, but not C2-domain, affects co-factor activity

Thus far, purified chimeras were characterized in absence of lipids while FVIII/FV C-domains are known to assemble on (negatively charged) membranes.[13-16] To study the function of the FVIII-YFP/FV chimeras, we examined complex formation with FIXa on phospholipid vesicles and their capability to convert the physiological substrate, FX. Membrane-binding was assessed by varying the phospholipid vesicle concentration at a fixed concentration of FIXa (16 nM). When using phospholipid vesicles containing 15% phosphatidylserine (PS), no reduction in FXa generation rates by FVIII-YFP/FV-C2 were observed in comparison with FVIII-YFP (Fig. 4A). A lower maximal FXa generation rate was seen with FVIII-YFP/FV-C1. However, both chimeras showed the same apparent affinity to membranes as FVIII-YFP (half-maximal response ~0.5 μM). The reduction in maximal response of FVIII-YFP/FV-C1 is due to reduced binding of FIXa, as was assessed by varying the FIXa concentration at a fixed phospholipid concentration (Fig. 4B). Therein, the FXa generation rates of FVIII-YFP/FV-C2 and FVIII-YFP are identical. FVIII-YFP/FV-C1, however, needed a concentration of at least 30 nM to reach the same level. Similar data were obtained in the presence of vesicles containing 5% PS, with the exception that the defect of the

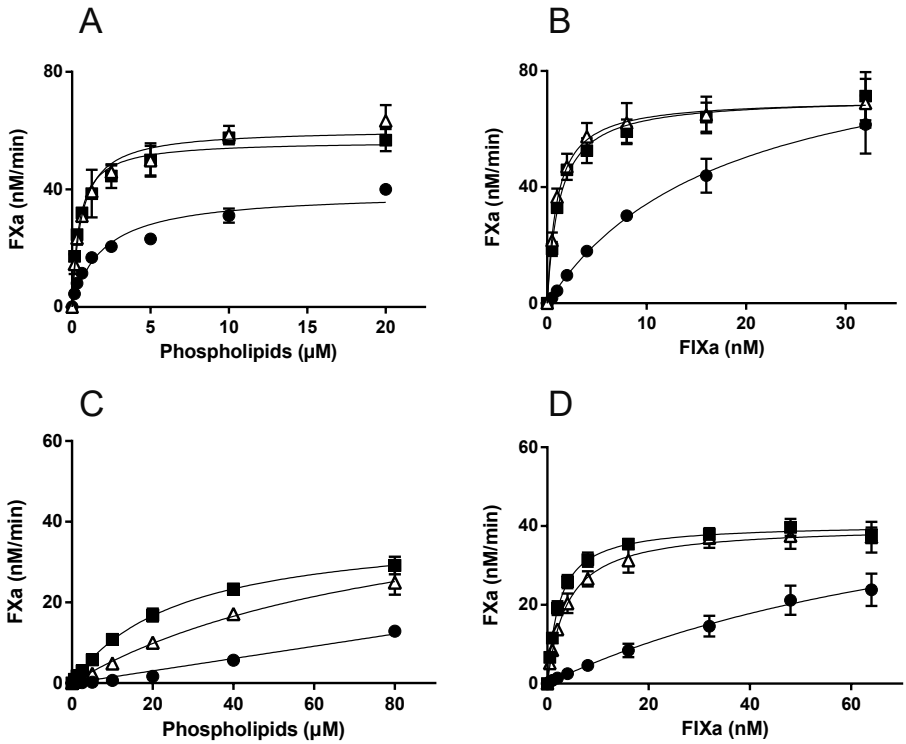


Figure 4: **FVIII cofactor activity of the FVIII-YFP chimeras.** FVIII-YFP (closed squares), FVIII-YFP/FV-C1 (closed circles) and FVIII-YFP/FV-C2 (open triangles) were examined for their capability to function as cofactor in the activation of FX. FX activation studies were performed in the presence of phospholipid vesicles containing 15% phosphatidylserine, 20% phosphatidylethanolamine, and 65% phosphatidylcholine (panels A and B) or 5% phosphatidylserine, 20% phosphatidylethanolamine, and 75% phosphatidylcholine panels (C and D) in a buffer containing 40 mM Tris-HCl (pH 7.8), 150 mM NaCl and 0.2% (w/v) bovine serum albumin. FIXa titrations (panels B and D) were performed by incubation of 25 μM of phospholipid vesicles mixed with 0-64 nM FIXa, 0.3 nM FVIII and 200 nM FX. Phospholipid titrations (panels A and C) were performed by incubation of 0-80 μM of phospholipid vesicles mixed with 16 nM FIXa, 0.3 nM FVIII and 200 nM FX. Reactions were initiated by addition of 1.5 mM $CaCl_2$ and 1 nM thrombin. Subsamples of the reaction mixture were taken at 30 sec intervals and analyzed for FXa using the substrate S-2765 containing the thrombin inhibitor I-2581.[21] Absorbance values were converted into molar concentrations using a standard curve of active-site titrated purified FXa.

FVIII-YFP/FV-C1 variant proved more prominent (Fig. 4D). At this low PS content the activity of FVIII-YFP/FV-C2 was also somewhat reduced at low lipid concentrations (Fig. 4C). These data demonstrate that in terms of co-factor activity the C2-domain is interchangeable with that of FV. However, swapping the C1-domain introduces a defect that involves assembly with FIXa.

Distorted C-domain pairing in chimeras containing FV C-domains

To examine the structural integrity of the chimeras, structural differences between FVIII-YFP and the chimeras were studied by lysine residue labeling with Tandem-Mass-Tags (TMTs). This approach consists of a pairwise comparison between proteins by labeling each with one of two isobaric labels (TMT-126 or TMT-127), which upon tandem mass spectrometry reveal their lysine-bound indicator mass of either 126 or 127 Da. This enables simultaneous identification and quantification of surface-exposed lysine residues.[22] We have used TMT-126 for FVIII-YFP and TMT-127 for the FVIII-YFP/FV chimeras and expressed relative incorporation of TMT labels in the ratio of TMT-127/TMT-126 per individual peptide. A ratio above 1 indicates increased lysine exposure in the FVIII-YFP/FV chimera (TMT-127 labeled) compared to the FVIII-YFP (TMT-126 labeled), and vice versa for ratios below 1.

Comparing the TMT-labeling incorporation of the FVIII-YFP/FV chimeras to FVIII-YFP reveals that most of the lysine residues' TMT-127/TMT-126 ratios were around 1 (*Supplementary Figure S3*). This indicates that accessibility for most of the lysine residues remained unchanged. However, some peptides showed a TMT-127/TMT-126 ratio > 1 in both chimeras (Fig. 5A). Most of these peptides contain lysine residues situated at the C2/A1-domain interface (Lys107, Lys123 and Lys127) (Fig. 5B). Peptides covering the C2-domain lysine residues Lys2239 and Lys2249 showed increased accessibility only in FVIII-YFP/FV-C1 (Fig. 5A). Due to the swapping of the C2-domain and thus the absence of TMT-127 labeled counterparts, FVIII-YFP/FV-C2 showed a near-zero TMT-127/TMT-126 ratio for these lysine residues. Apart from lysine residues with increased accessibility, FVIII-YFP/FV-C2 also displayed lysine residues in the C1-domain with decreased accessibility. These are Lys2020, Lys2110 and/or Lys2111. (Fig. 5A, B and *Supplementary Figure S3*).

Using TMT labeling, the lysine exposure of the swapped C-domains themselves could also be investigated in comparison with their native conformation in FV). This was done by labeling of FV with TMT-126 and comparing it to the chimeras labeled with TMT-127. In this way only the swapped C-domains are examined. Again, most peptides displayed an unchanged lysine exposure (ratios are approx. 1). However, one C1-domain peptide containing two lysine residues directed towards the opposing C2-domain in the FV model had an average TMT-127/TMT-126 ratio of 3 (Lys1941 and Lys1954) (Fig. 6A, C). Within the swapped C2-domain, two peptides indicated an increased accessibility of its lysine residues: one peptide with lysine residues directed towards to the C1-domain (Lys2157 and Lys2161) and one peptide containing a lysine residue directed towards the A1-domain (Lys2137) (Fig. 6B, C). Compared to FV, these lysine residues are more accessible for TMT labeling in the chimeras. Thus, taken together the comparisons between FVIII-YFP, FV and the chimeras, the C1/C2- and C2/A1-domain interfaces become exposed which suggests

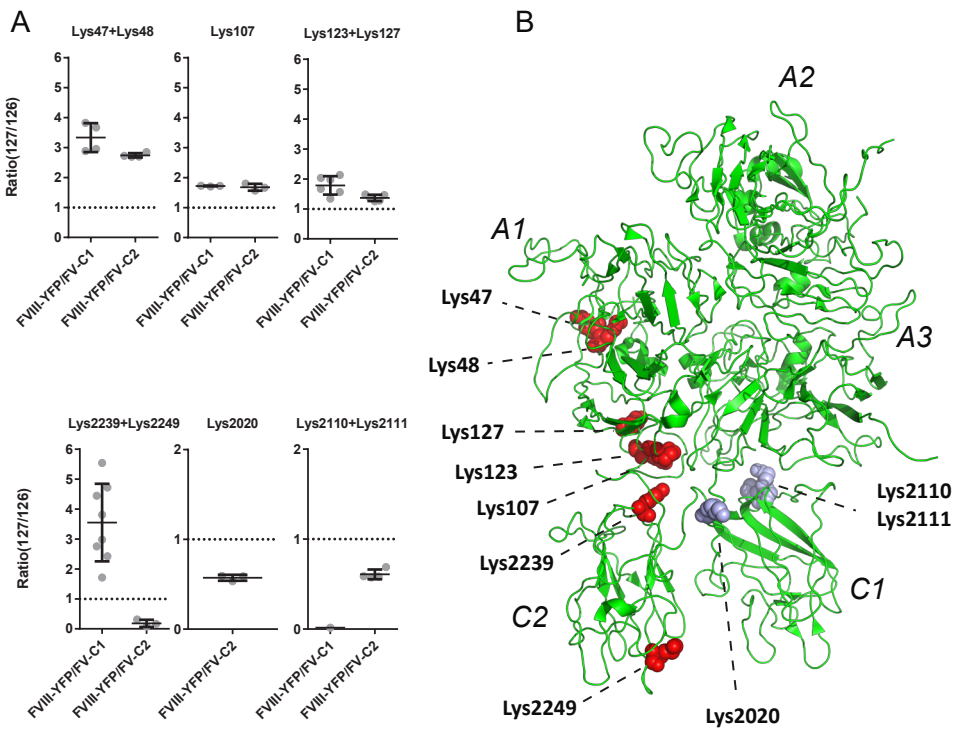


Figure 5: **Comparison of FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2 with FVIII-YFP by TMT-labeling.** FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2 were labeled with TMT-127 and FVIII-YFP with TMT-126 to assess differences in lysine residue exposure. Labeled FVIII variants were then processed into peptides and measured for TMT incorporation. The lysine residues with a TMT-127/TMT-126 ratio different from 1 are: Lys47+Lys48, Lys107, Lys123+Lys127, Lys2020, Lys2110+Lys2111 and Lys2239+2249. **(A)** The peptides containing the same lysine residue(s) and their TMT-127/TMT-126 ratio per chimera are shown in the different panels. The mean TMT-127/TMT-126 ratio ±SD is given by the black lines; peptides and their resolved TMT-127/TMT-126 ratio are represented by gray dots. A TMT-127/TMT-126 ratio above or below 1 (dashed line) indicates increased or decreased incorporation of TMT-127 and therefore altered accessibility of those lysine residues in the chimera compared to FVIII-YFP. Peptides covering lysine residues 2239 and 2249 show a TMT-127/TMT-126 ratio > 1 for chimera FVIII-YFP/FV-C1. In the FVIII-YFP/FV-C2 chimera these lysine residues have a near zero ratio because the FVIII C2-domain peptides are not present in the FVIII-YFP/FV-C2 chimera (only TMT-126 labeled FVIII-YFP lysine residues could be detected). Software analysis considers a TMT-127/TMT-126 ratio of zero erroneous and therefore may produce ratios with a minimum value of 0.01. **(B)** Lysine residues with increased accessibility for TMT labeling are indicated in the FVIII crystal structure (2R7E)[30] as red spheres, whereas lysine residues with reduced reactivity are indicated in blue.

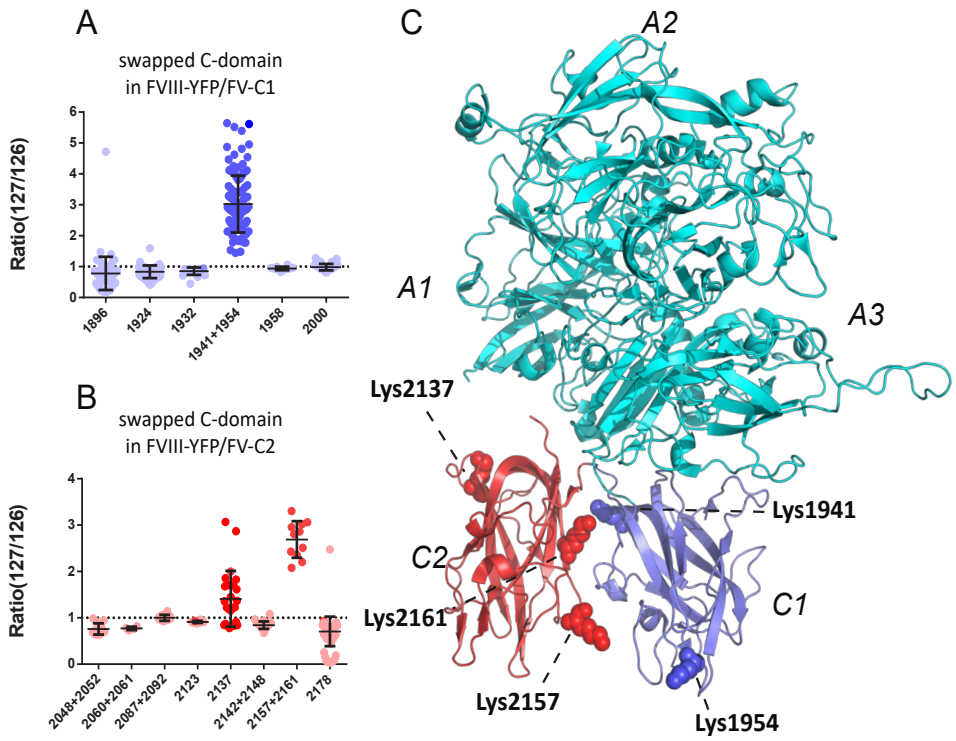


Figure 6: **Comparison of FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2 with FV by TMT-labeling.** FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2 were labeled with TMT-127, and FV with TMT-126. The mean TMT-127/TMT-126 ratio ± SD (black lines) is given for peptides containing the same lysine residues from **(A)** the C1-domain in blue and **(B)** the C2-domain in red of FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2, respectively. Discrete peptides and their measured TMT-127/TMT-126 ratio are depicted by a dot. Lysine residues with a TMT-127/TMT-126 ratio above 1 (dashed line) are Lys1941+Lys1954, Lys2137 and Lys2157+Lys2161. Such a TMT-127/TMT-126 ratio indicates increased incorporation of TMT-127 and therefore increased accessibility of those lysine residues in the chimera compared to FV wild-type. **(C)** The lysine residues with TMT-127/TMT-126 ratios above 1 are indicated in a model of FVa (1Y61)[32] with spheres. The C1-domain is indicated in blue and the C2-domain in red.

that pairing of the C-domains is altered in both FVIII-YFP/FV chimeras.

DISCUSSION

The site of FVIII biosynthesis has remained a matter of debate. It was previously established that FVIII overexpression in blood outgrowth endothelial cells (BOECs) leads to co-storage of FVIII with VWF in WPBs.[9,10] Relevance of this model system has recently been established by findings that FVIII is expressed in endothelial cells.[6-8] However, the mechanism behind FVIII trafficking and secretion remains poorly understood. Here we studied the expression of FVIII, FV and FVIII/FV C-domain chimeras in BOECs. In contrast to FV, FVIII was solely found in WPBs (Fig. 1). This implies a specific FVIII mechanism, which appears to be affected once one of the two FVIII C-domain is swapped with FV (Fig. 2). Particularly, C1-domain swapping appeared destructive, although FVIII-YFP/FV-C2 also displayed a trafficking defect as well. Thus, both C-domains appear to contribute. Remarkably, C-domains that tightly bind to phospholipids tend to occur in a tandem pair as observed in FVIII, FV, lactadherin and developmental-endothelial-locus 1 (del-1).[27] Next to a tandem or multimeric organisation, secretion of these strong lipid-binders tends to be regulated. For instance, lactadherin released by several cell types is associated with confined membrane vesicles (exosomes).[28,29] As reported by others, deletion of one of the two lactadherin C-domains results in a loss of the exosome-mediated secretion.[29] This suggests that trafficking of lactadherin requires an intact C-domain tandem pair, which might also be required for FVIII.

Within both FVIII-YFP/FV chimeras, TMT footprinting revealed more accessibility between the A1/C2- and C1/C2-domain interface implying a loose C-domain tandem pair with increased C2-domain mobility (Fig. 5 and 6). This extends the notion that in FVIII the C2-domain has limited interdomain contacts.[30] In two available FVIII crystal structures,[30,31] the C2-domain Lys2239 interaction with a nearby glutamic acid (Glu122 in the A1-domain) seems to be the only evident electrostatic interaction (*Supplementary Figure S4*). A swap of the C2-domain may affect this interaction even though Lys2239 is conserved in FV. The swapped Lys2239 could not be detected, however, Glu122 neighbouring lysine residues Lys123 and Lys127 could be measured and displayed an increased TMT-127/TMT-126 ratio (Fig. 5). Strikingly, when swapping the C1-domain, Lys2239 could be resolved and also displayed increased labelling despite its remote location. Perhaps, introduction of a FV C-domain presents an unfavourable interaction in the C1/C2-domain interface which in turn affects the C2/A1-domain interface. By one-sided insertion of a FV C-domain in FVIII, interactions would be disrupted due to loss of the interacting counterparts.

Indeed, Lys1941 and Lys2161 display an increased accessibility in the chimeras while in FV they most likely interact with residues on their opposite C-domain Phe2163 and Glu2034 (*Supplementary Figure S4*).[32] The A3/C1-domain interface could hardly be probed because it contains a limited number of lysine residues. Nonetheless three of these residues, at positions 2010, 2110 and 2111 in the C1-domain, displayed reduced surface exposure in the FVIII-YFP/FV-C2 chimera (Fig. 5 and *Supplementary Fig. S3*). The apparent protection of these lysine residues suggests that the top of the C1-domain associates more tightly with the A3-domain in this chimera, thus maintaining the A3/C1-domain interface undisrupted.

An overall limited disruption of the domain organisation is reflected in the activity of the FVIII/FV-C2 chimera. The FVIII/FV-C2 chimera, despite its more accessible C1/C2- and C2/A1-domain interfaces, displays normal FIXa binding and FX activation. Moreover, swapping the entire C2-domain within FVIII does not result in any functional loss (Fig. 4B, D). Interestingly, as reported by others, deletion of the C2-domain results in a 4-fold reduction in FIXa affinity.[33] Apparently, the presence of the C1-domain alone is insufficient for full co-factor function. However, the C2-domain of FV fully compensates for this defect (Fig. 4B). This apparent paradox can be explained by a complementary role for the C2-domain in its FIXa-binding conformation. Unlike the C2-domain, the C1-domain contributes directly to FIXa binding. This is in agreement with the observation of Wakabayashi *et al*., who noted that a FVIII variant with the C1-domain replaced by a second C2-domain exhibits an approximately 9-fold reduced affinity for FIXa.[34]

The FVIII/FV-C1 chimera has a VWF-binding defect that is more severe than that of FVIII lacking the sulfated tyrosine on position 1680 (FVIII-Y1680F).[11,12] This is even more prominent at pH 5.5 (Fig. 3 and *Supplementary Fig. S2*). This implies that apart from the sulfated Tyr1680, the FVIII C1-domain is essential for the VWF/FVIII complex formation. This is in agreement with previous reports that mutations in the C1-domain can interfere with VWF-binding,[5,12] although this does not exclude an additional contribution of the C2-domain.[4] A direct C1-domain interaction with VWF is further supported by recent studies using hydrogen-deuterium exchange and electron microscopy.[35,36]

Previously, we analyzed a variety of FVIII mutations that are associated with reduced interaction with VWF, but normal trafficking to WBPs.[12] The fact that these variants retained some residual VWF binding at pH 5.5 led us to speculate that this might be sufficient for trafficking to WPBs.[12] In this respect the FVIII/FV-C1 chimera is more defective than the haemophilia A causing mutations which we studied. If FVIII storage with VWF would be exclusively driven by VWF-binding, this might explain the trafficking defect for FVIII/FV-C1 (Fig. 2B). Surprisingly, although the VWF-binding of FVIII/FV-C2 is close to normal, this chimera also displays reduced WPB storage and

intracellular accumulation in BOECs (Fig. 2C). This supports the conclusion that FVIII trafficking to WPBs requires the tandem C-domain pair. This would be compatible with our previous data, because none of the haemophilia A causing mutations that we analysed can be expected to disrupt the C-domain pair [12]

One limitation of lentiviral BOEC transduction is the variability in expression which precludes obtaining quantitative information to directly compare the extent of WPB trafficking with VWF-binding affinity. As we showed previously, this issue can be addressed by subcellular fractionation, provided that expression and transduction efficiency are in the same order of magnitude.[10,12] However, this proved unfeasible for the present set of FVIII/FV chimeras. Nevertheless, it seems evident that the intracellular YFP-staining for the FVIII/FV-C2 chimera, despite its high affinity for VWF, is different from that of FVIII-YFP or BDD-FVIII (Fig. 1A and 2A, C). Likewise, the FVIII/FV-C1 chimera, which displays no appreciable VWF interaction, could not be visualized intracellularly (Fig. 2B). Despite these trafficking defects in BOECs, the chimeras could be expressed in HEK293-cells, and after purification displayed structural integrity and functionality (Fig. 4-6 and *Supplementary Fig. S3*). This might suggest that besides VWF-binding another yet unknown mechanism could play a role in endothelial FVIII storage and secretion. A role for the A1-domain could therein not be excluded as the chimeras, compared to FVIII-YFP, display a structural change in the A1-domain as well (Lys47+Lys48, Fig. 5). Whether or not these A1-domain residues contribute to maintaining the C-domain tandem organisation remains an open question.

Our finding that the C2-domain of FVIII can be replaced by that of FV without compromising FVIII activity may have translational implications. It has been well established that haemophilia A patients with FVIII inhibitors often have antibodies against the C2-domain.[37] It seems conceivable that such inhibitors may be bypassed by FVIII containing the FV C2-domain. Because swapping the C1 domain eliminates VWF binding and affects FIXa interaction, the potential of FVIII containing the FV C1-domain for bypassing C1-domain directed inhibitors seems less evident.

Acknowledgements

The authors wish to thank J.M. Koornneef for providing purified factor V. This study was supported by the Landsteiner Foundation for Blood Transfusion Research (LSBR).

REFERENCES

1. Lenting PJ, van Mourik JA, Mertens K. The life cycle of coagulation factor VIII in view of its structure and function. *Blood.* 1998 **92** 3983-3996.
2. Thompson AR. Structure and function of the factor VIII gene and protein. *Semin Thromb Hemost.* 2003 **29** 11-22.
3. Leyte A, van Schijndel HB, Niehrs C, *et al*. Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. *J Biol Chem.* 1991 **266** 740-746.
4. Saenko EL, Scandella D. The acidic region of the factor VIII light chain and the C2 domain together form the high affinity binding site for von Willebrand factor. *J Biol Chem.* 1997 **272** 18007-18014.
5. Jacquemin M, Lavend'homme R, Benhida A, *et al*. A novel cause of mild/moderate hemophilia A: mutations scattered in the factor VIII C1 domain reduce factor VIII binding to on Willebrand factor. *Blood.* 2000 **96** 958-965.
6. Fahs SA, Hille MT, Shi Q, Weiler H, Montgomery RR. A conditional knockout mouse model reveals endothelial cells as the principal and possibly exclusive source of plasma factor VIII. *Blood.* 2014 **123** 3706-3713.
7. Everett LA, Cleuren AC, Khoriaty RN, Ginsburg D. Murine coagulation factor VIII is synthesized in endothelial cells. *Blood.* 2014 **123** 3697-3705.
8. Shahani T, Covens K, Lavend'homme R, *et al*. Human liver sinusoidal endothelial cells but not hepatocytes contain factor VIII. *J Thromb Haemost.* 2014 **12** 36-42.
9. Rosenberg JB, Foster PA, Kaufman RJ, *et al*. Intracellular trafficking of factor VIII to von Willebrand factor storage granules. *J Clin Invest.* 1998 **101** 613-624.
10. Van den Biggelaar M, Bouwens EAM, Kootstra NA, Hebbel RP, Voorberg J, Mertens K. Storage and regulated secretion of factor VIII in blood outgrowth endothelial cells. *Haematologica.* 2009 **94** 670-678.
11. Van den Biggelaar M, Bierings R, Storm G, Voorberg J, Mertens K. Requirements for cellular co-trafficking of factor VIII and von Willebrand factor to Weibel-Palade bodies. *J Thromb Haemost.* 2007 **5** 2235-2242.
12. Van den Biggelaar M, Bouwens EA, Voorberg J, Mertens K. Storage of factor VIII variants with impaired von Willebrand factor binding in Weibel-Palade bodies in endothelial cells. *PLoS One.* 2011 **6** e24163.
13. Lu J, Pipe SW, Miao H, Jacquemin M, Gilbert GE. A membrane-interactive surface on the factor VIII C1 domain cooperates with the C2 domain for cofactor function. *Blood.* 2011 **117** 3181-3189.
14. Gilbert GE, Novakovic VA, Kaufman RJ, Miao H, Pipe SW. Conservative mutations in the C2 domains of factor VIII and factor V alter phospholipid binding and cofactor activity. *Blood.* 2012 **120** 1923-1932.

15. Gilbert GE, Kaufman RJ, Arena AA, Miao H, Pipe SW. Four hydrophobic amino acids of the factor VIII C2 domain are constituents of both the membrane-binding and von Willebrand factor-binding motifs. *J Biol Chem.* 2002 **277** 6374-6381.
16. Peng W, Quinn-Allen MA, Kane WH. Mutation of hydrophobic residues in the factor Va C1 and C2 domains blocks membrane-dependent prothrombin activation. *J Thromb Haemost.* 2005 **3** 351-354.
17. Adams TE, Hockin MF, Mann KG, Everse SJ. The crystal structure of activated protein C-inactivated bovine factor Va: Implications for cofactor function. *Proc Natl Acad Sci U S A.* 2004 **101** 8918-8923.
18. Bos MH, Camire RM. Blood coagulation factors V and VIII: Molecular mechanisms of procofactor activation. *J Coagul Disord.* 2010 **2** 19-27.
19. Bos MHA, Meijerman DWE, van der Zwaan C, Mertens K. Does activated protein C-resistant factor V contribute to thrombin generation in hemophilic plasma? *J Thromb Haemost.* 2005 **3** 522-530.
20. Meems H, van den Biggelaar M, Rondaij M, van der Zwaan C, Mertens K, Meijer AB. C1 domain residues Lys 2092 and Phe 2093 are of major importance for the endocytic uptake of coagulation factor VIII. *Int J Biochem Cell Biol.* 2011 **43** 1114-1121.
21. Bloem E, van den Biggelaar M, Wroblewska A, Voorberg J, *et al*. Factor VIII C1 domain spikes 2092-2093 and 2158-2159 comprise regions that modulate cofactor function and cellular uptake. *J Biol Chem.* 2013 **288** 29670-29679.
22. Bloem E, Ebberink EH, van den Biggelaar M, van der Zwaan C, Mertens K, Meijer AB. A novel chemical footprinting approach identifies critical lysine residues involved in the binding of receptor-associated protein to cluster II of LDL receptor-related protein. *Biochem J.* 2015 **468** 65-72
23. Bouwens EA, Mourik MJ, van den Biggelaar M, *et al*. Factor VIII alters tubular organization and functional properties of von Willebrand factor stored in Weibel-Palade bodies. *Blood.* 2011 **118** 5947-5956.
24. Van den Biggelaar M, Meijer AB, Voorberg J, Mertens K. Intracellular cotrafficking of factor VIII and von Willebrand factor type 2N variants to storage organelles. *Blood.* 2009 **113** 3102-3109.
25. Van den Biggelaar M, Madsen JJ, Faber JH, *et al*. Factor VIII interacts with the endocytic receptor low-density lipoprotein receptor-related protein 1 via an extended surface comprising "hot-spot" lysine residues. *J Biol Chem.* 2015 **290** 16463-76.
26. Wu MM, Grabe M, Adams S, Tsien RY, Moore HP, Machen TE. Mechanisms of pH regulation in the regulated secretory pathway. *J Biol Chem.* 2001 **276** 33027-33035.
27. Vogel W. Discoidin domain receptors: structural relations and functional

implications. *FASEB J*. 1999 **13** S77-S82.
28. Véron P, Segura E, Sugano G, Amigorena S, Théry C. Accumulation of MFG-E8/lactadherin on exosomes from immature dendritic cells. *Blood Cells Mol Dis.* 2005 **35** 81-88.
29. Oshima K, Aoki N, Kato T, Kitajima K, Matsuda T. Secretion of a peripheral membrane protein, MFG-E8, as a complex with membrane vesicles. *Eur J Biochem.* 2002 **269** 1209-1218.
30. Shen BW, Spiegel PC, Chang C-H, *et al*. The tertiary structure and domain organization of coagulation factor VIII. *Blood.* 2008 **111** 1240-1247.
31. 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.
32. Orban T, Kalafatis M, Gogonea V. Completed three-dimensional model of human coagulation factor Va. Molecular dynamics simulations and structural analyses. *Biochemistry.* 2005 **44** 13082-13090.
33. Wakabayashi H, Griffiths AE, Fay PJ. Factor VIII lacking the C2 domain retains cofactor activity in vitro. *J Biol Chem.* 2010 **285** 25176-25184.
34. Wakabayashi H, Fay PJ. Replacing the factor VIII C1 domain with a second C2 domain reduces factor VIII stability and affinity for factor IXa. *J Biol Chem.* 2013 **288** 31289-31297.
35. Chiu PL, Bou-Assaf GM, Chhabra ES, *et al*. Mapping the interaction between factor VIII and von Willebrand factor by electron microscopy and mass spectrometry. *Blood.* 2015 **126** 935-938.
36. Yee A, Oleskie AN, Dosey AM, *et al*. Visualization of an N-terminal fragment of von Willebrand factor in complex with factor VIII. *Blood.* 2015 **126** 939-942.
37. Meeks SL, Healey JF, Parker ET, Barrow RT, Lollar P. Antihuman factor VIII C2 domain antibodies in hemophilia A mice recognize a functionally complex continuous spectrum of epitopes dominated by inhibitors of factor VIII activation. *Blood.* 2007 **110** 4234-4242.

SUPPLEMENTARY SECTION

MATERIAL AND METHODS

Materials

All chemicals used were of analytical grade. Fetal calf serum was from Hyclone (Logan, UT, USA). One shot *Stbl3* chemically competent cells, trypsin, and DMEM (4.5 g/L glucose) were obtained from Invitrogen (Breda, the Netherlands). DNA modifying enzymes were from NEB (Ipswich, MA, USA). EGM-2 bullet kit, penicillin and streptomycin were from Lonza (Walkersville, MD, USA). Collagen type 1 rat tail was from BD Biosciences (Uppsala, Sweden). Peroxidase-labeled polyclonal rabbit anti-human VWF antibody was supplied by DAKO (Glostrup, Denmark). Tissue culture flasks, multidishes and microtiterplates (Maxisorp) were from Nunc (Roskilde, Denmark). TMTs, hydroxylamine and chymotrypsin were from Thermo Scientific (Breda, the Netherlands). The chromogenic substrate S-2765 containing the thrombin inhibitor I-2581 was from the Coatest FVIII kit (Chromogenix-Instrumentation Laboratory, Milano, Italy).

Immunofluorescence microscopy of FVIII and FV

BOECs were grown on collagen-coated 1 cm-diameter glass coverslips. Confluent cells were fixed with 3.7% PFA in PBS for 15 minutes at room temperature. After fixation, the cells were labeled with monoclonal antibody CLB-RAg20[1] to detect VWF, followed by incubation with Alexa-633 conjugated secondary antibody and FITC-labeled EL-14[2] (directed against FVIII C2-domain), FITC-labeled KM33[3] (against FVIII C1-domain) or FITC-labeled CLB-FV-4 (directed against FV light chain).[4] Cells were embedded in 4-88 Mowiol and stored at 4 °C until analysis. Z-stacks (0.4-µm intervals) were taken with confocal laser scanning microscopy using a Zeiss LSM510 equipped with Plan NeoFluar 63x/1.4 Oil objective (Carl Zeiss, Heidelberg, Germany). Images were processed with Zeiss LSM510 version 4.0 software and LSM image browser (Carl Zeiss, Heidelberg, Germany). Quantification of FVIII variants and FV antigen levels in the conditioned medium by enzyme-linked immunosorbent assay (ELISA) was essentially done as described elsewhere,[5,6] except that for determination of FV antigen, the immobilized monoclonal antibody CLB-FV-44 was used and purified FV with known concentration served as reference.

Construction of FVIII-YFP/FV chimera pcDNA3.1(+) vectors

For construction of FVIII containing a FV C1- or C2-domain, FV C-domains were amplified from pcDNA3.1(+) encoding B-domain deleted FV 811-1491[6] using the following primers for the FV C1-domain:

5′-ATAACCGGTGGACTAAGCACTGGTATC-3′ (sense) and
5′-TATTGGCGCGCCACCATTTACCTCACAACCTTG-3′(antisense).

For the FV C2-domain the following primers were used:

5′-ATAACGGTGGATGTTCCACACCCCTG-3′ (sense) and
5′-TATTGGCGCGCCACCGTAAATATCACAGCCAAA-3′ (antisense).

PCR products were ligated in pGEM-T Easy vector (Promega) and digested with *AgeI* and *AscI*. The FVIII C1- (ranging from Cys2021 to Asp2172) or C2-domain (ranging from Ser2173 to the C-terminal end) in pcDNA3.1(+) FVIIIYFP[7] were replaced for *AgeI-AscI* restriction sites by Quick Change mutagenesis using the following primers for the C1-domain (restriction sites are underlined):

5′-CTGGTGTACAGCAATAAG<u>ACCGGTGGCGCGCCA</u>AGTTGCAGCATGCC
ATTG-3′ (sense) and
5′-CAATGGCATGCTGCAACT<u>TGGCGCACCGGT</u>CTTATTGCTGTAC
ACCAG-3′ (antisense).

Constructs were digested with *AgeI-AscI* and ligated with corresponding FV C-domain fragments. *AgeI* and *AscI* restriction sites bordering the C-domains were removed by Quick Change mutagenesis using appropriate primers. Coding sequences of the constructs were analyzed by sequencing. After sequencing we found that the first 5 amino acids of the swapped C1-domain were missing in the FVIII-YFP/FV-C1 variant. These were added using Quick Change mutagenesis with the following primers:

5′-CTGGTGTACAGCAATAAGTGTAGGATGCCAATGGGACTAAGCACTGGT
ATC-3′ (sense) and
5′-GATACCAGTGCTTAGTCCCATTGGCATCCTACACTTATTGCTGTACAC
CAG-3′ (antisense).

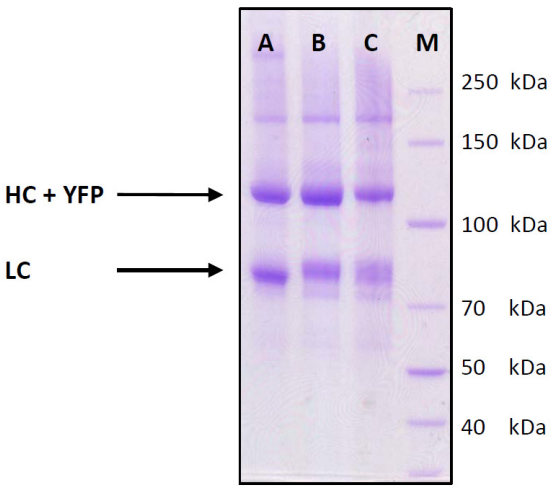
Construction of lentiviral vectors

Lentiviral plasmids encoding B-domain deleted FVIII (FVIII) and FVIII/FV chimeras were created by replacing fragment *NheI-XhoI* from pLV-CAG-FVIIIGFP vector[5] for the corresponding fragment of the pcDNA3.1(+) plasmids. To clone BDD-FV in lentiviral vector, pLV-CAG-FVIIIGFP vector was first digested with *NheI* followed by partial digestion with *NotI*. Briefly, *NheI*-digested construct was mixed on ice with serial dilutions of *NotI*, incubated for 15 minutes at 37 °C and heated to 65 °C for 20 minutes to inactivate the restriction enzyme. The fragment corresponding to lentiviral vector without insert was isolated from gel and ligated with the *NheI-NotI* fragment of FV in pcDNA3.1(+). Production of lentivirus, BOEC isolation and transduction with lentivirus have been described before.[5]

TMT-labeling of FVIII-YFP variants and mass spectrometry analysis

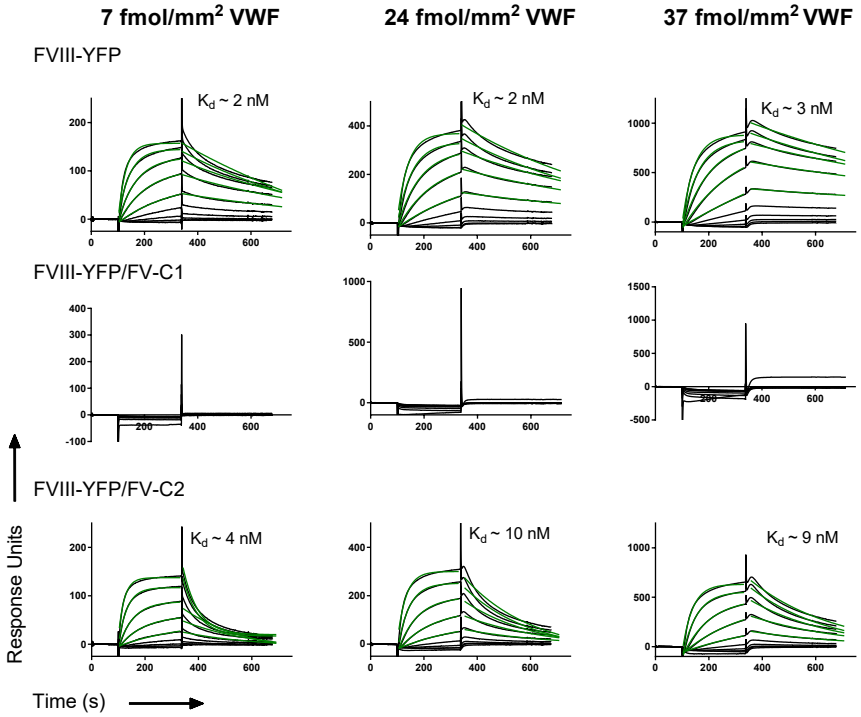
TMT-labeling was done as essentially described elsewhere.[8] Briefly, labeling was initiated by addition of 2.5 mM of TMT labels to 140 nM of the chimeric FVIII-YFP/FV protein or 140 nM of the normal FVIII-YFP protein in buffer containing 20 mM HEPES, 5 mM $CaCl_2$ and 175 mM NaCl and incubated at 25 °C (Eppendorf Thermomixer comfort, 600 rpm). In our experiments, TMT-127 was used for the chimeric protein and TMT-126 was used for the FVIII-YFP or FV. The labeling reaction was quenched after 15 minutes by addition of 500 mM of hydroxylamine. TMT-labeled wild-type protein was pooled in a 1:1 molar ratio with either TMT-labeled FVIII-YFP/FV-C1 or FVIII-YFP/FV-C2. The pooled protein mixture was reduced and acetylated as described previously.[8] Peptides were obtained after digestion by chymotrypsin with subsequent desalting using a C18 Ziptip (Millipore, Amsterdam, the Netherlands), both performed according to the manufacturer's instructions. Peptides were separated by reverse phase chromatography and sprayed into a Thermo Scientific LTQ OrbitrapXL mass spectrometer. Peptides from the chimeras and FVIII-YFP were separated on a C18 column with 0.5% (v/v) acetic acid using two different gradients from 0% up to 60% (v/v) of acetonitrile: a (short) 60 minute gradient and a (long) 160 minute gradient both with a wash peak of 98% (v/v) acetonitrile for 5 minutes. Peptides acquired from the chimeras and FV wild-type were separated using the long 160 minute gradient. Peptide identification and TMT quantification by mass spectrometric analysis was done as described elsewhere.[8]

SUPPLEMENTARY FIGURE S1

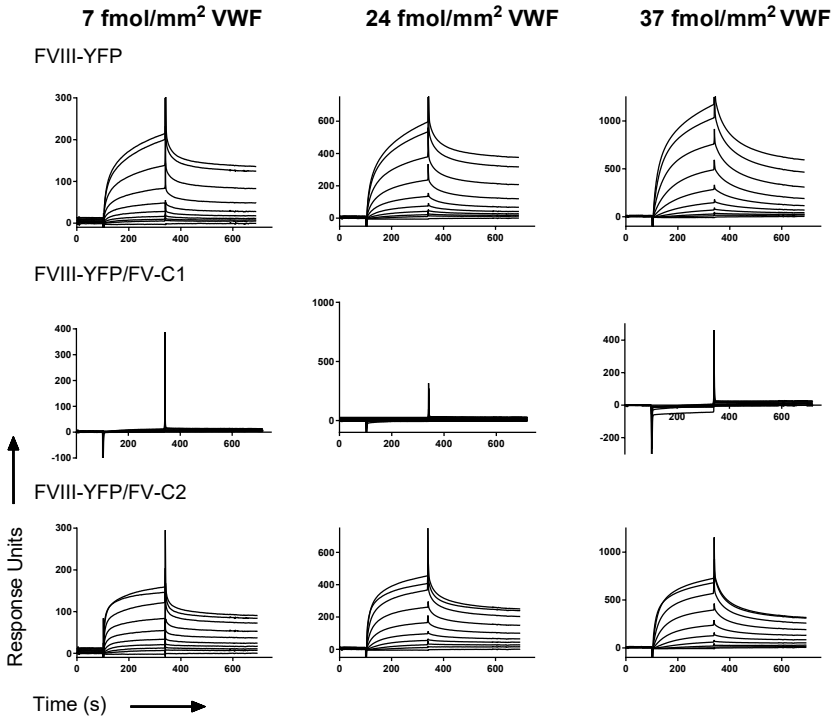


Supplementary Figure S1: **SDS PAGE analysis of purified chimeras. (A)** Purified FVIII-YFP, **(B)** FVIII-YFP/FV-C1 and **(C)** FVIII-YFP/FV-C2 were run on a 4-12% gradient SDS PAGE (Invitrogen). Indicated are the YFP-carrying heavy chain domain (HC+YFP) of approx. 110 kDa and light chain (LC) of approx. 80 kDa as well as a minor single chain component. Masses were inferred using a PageRuler Broad Range (Thermo Scientific) marker (M).

SUPPLEMENTARY FIGURE S2

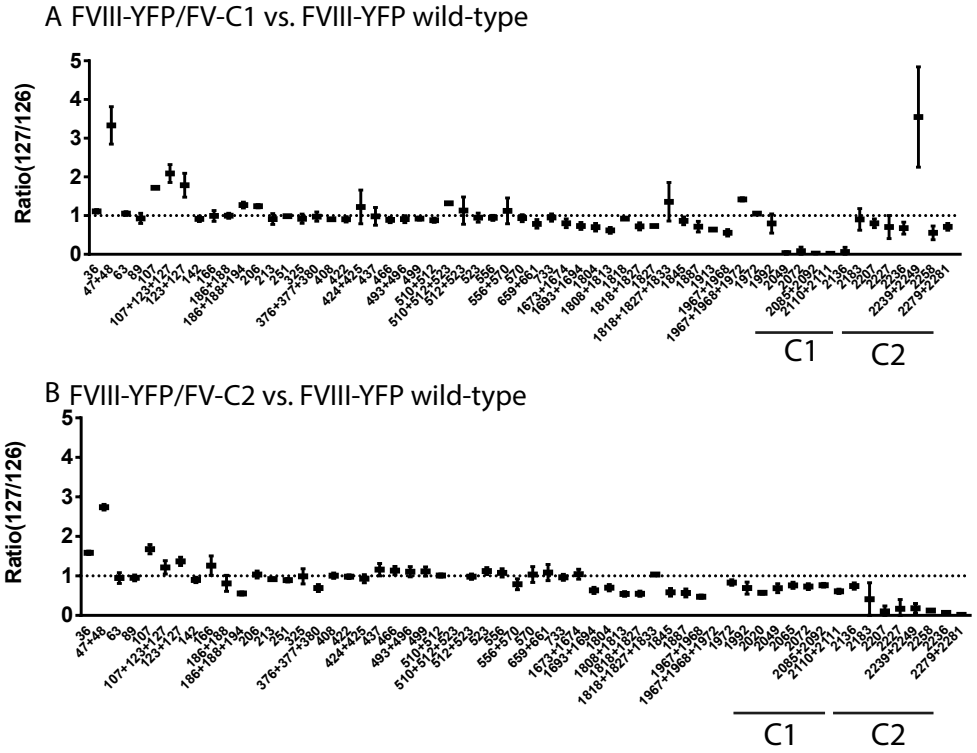


(figure legend on next page)



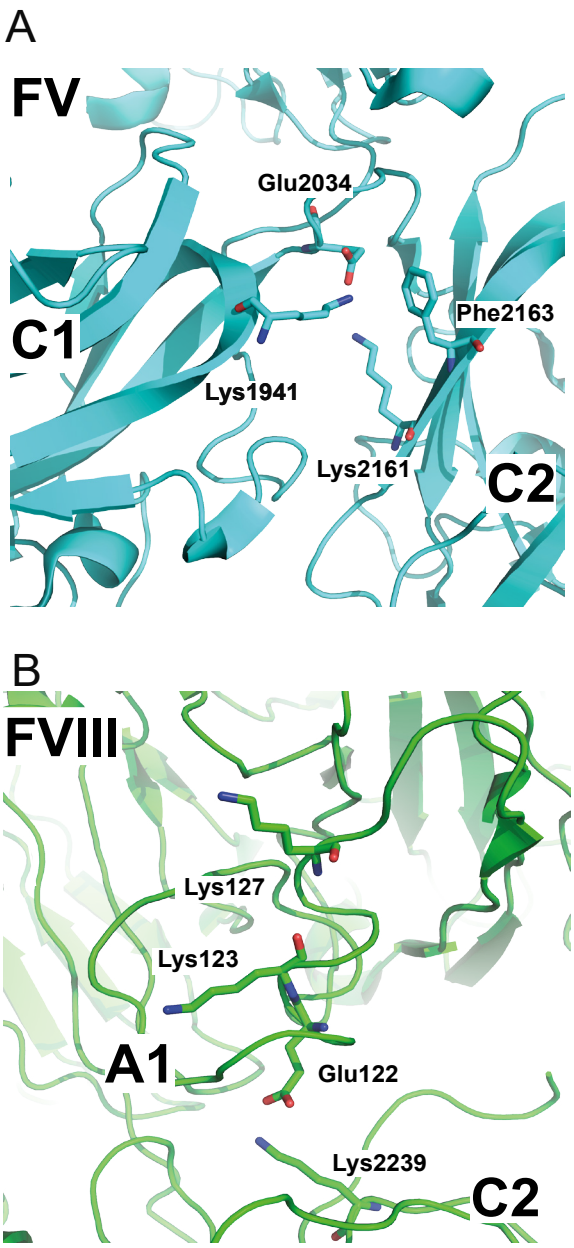
Supplementary Figure S2: **Interaction of FVIII-YFP variants with immobilized VWF.** Displayed are the sensorgrams acquired during the SPR experiments done at **(A)** pH 7.4 and **(B)** pH 5.5. FVIII-YFP variants with different concentrations ranging from 0.23 to 60 nM were passed over a chip with different VWF densities (7 fmol/mm^2, 24 fmol/mm^2 and 37 fmol/mm^2). Spikes are due to minor differences in buffer composition between samples and running buffer. Ignoring these imperfections, sensorgrams of FVIII-YFP and FVIII-YFP/FV-C2 were fitted using a Langmuir 1:1 binding model.[9] The association phase was fitted using a k_{off} value obtained from a separate dissociation phase fitting (also using a Langmuir 1:1 binding model).[9] The obtained K_d values of the binding curves at pH 7.4 are an average of the K_d values obtained from the fitted traces displayed in green solid curves. At pH 5.5 the dissociation phase displays partially irreversible binding which precluded calculation of K_d values.

SUPPLEMENTARY FIGURE S3



Supplementary Figure S3: **Overview of TMT-labeled FVIII-YFP/FV-C1 and FVIII-YFP/FV-C2 in comparison with FVIII-YFP.** The mean TMT-127/TMT-126 ratio with ± SD (black lines) is given for peptides containing the same lysine residues in **(A)** FVIII-YFP/FV-C1 and **(B)** FVIII-YFP/FV-C2 compared to FVIII-YFP. Most of the TMT-127/TMT-126 ratios are around 1, indicating unchanged lysine accessibility. The TMT-127/TMT126 ratios for lysine residues in the C1-domain or C2-domain are near zero for the respective experiments were either the FVIII-YFP/FV-C1 chimera or the FVIII-YFP/FV-C2 chimera was compared to FVIII-YFP. This could be explained because peptides containing lysines of FVIII-YFP are not present in the chimeras' swapped C-domains. Therefore, only FVIII-YFP TMT-126 labeled lysine residues could be detected.

SUPPLEMENTARY FIGURE S4



Supplementary Figure S4: **Potentially disrupted C1/C2 domain and A1/C2 domain interfaces. (A)** Within FV, two lysine residues Lys1941 and Lys2161 on the opposite C-domain could interact with Glu2034 (salt-bridge) or Phe2163 (cation-π). Swapping of a C-domain would leave half the interacting residues disengaged. **(B)** Within FVIII, Lys2239 appears to interact with Glu122 on the A1-domain. Disruption of this interaction could also have its effect on the nearby located lysine residues Lys123 and Lys127. Images were created with PyMol (Schrödinger, Cambridge, MA, USA).

SUPPLEMENTARY REFERENCES

1. Stel HV, Sakariassen KS, Scholte BJ, *et al*. Characterization of 25 monoclonal antibodies to factor VIII-von Willebrand factor: relationship between ristocetin-induced platelet aggregation and platelet adherence to subendothelium. *Blood.* 1984 **63** 1408-1415.
2. Van den Biggelaar M, Meijer AB, Voorberg J, Mertens K. Intracellular cotrafficking of factor VIII and von Willebrand factor type 2N variants to storage organelles. *Blood.* 2009 **113** 3102-3109.
3. Van den Brink EN, Turenhout EA, Bovenschen N, *et al*. Multiple VH genes are used to assemble human antibodies directed toward the A3-C1 domains of factor VIII. *Blood.* 2001 **97** 966-972.
4. Bos M: Regulatory mechanisms in the activation of blood coagulation factors V and VIII. PhD thesis, University of Utrecht, the Netherlands, 2006
5. Van den Biggelaar M, Bouwens EAM, Kootstra NA, Hebbel RP, Voorberg J, Mertens K. Storage and regulated secretion of factor VIII in blood outgrowth endothelial cells. *Haematologica.* 2009 **94** 670-678.
6. Bos MHA, Meijerman DWE, van der Zwaan C, Mertens K. Does activated protein C-resistant factor V contribute to thrombin generation in hemophilic plasma? *J Thromb Haemost.* 2005 **3** 522-530.
7. Van den Biggelaar M, Bierings R, Storm G, Voorberg J, Mertens K. Requirements for cellular co-trafficking of factor VIII and von Willebrand factor to Weibel-Palade bodies. *J Thromb Haemost.* 2007 **5** 2235-2242.
8. Bloem E, Ebberink EH, van den Biggelaar M, van der Zwaan C, Mertens K, Meijer AB. A novel chemical footprinting approach identifies critical lysine residues involved in the binding of receptor-associated protein to cluster II of LDL receptor-related protein. *Biochem J.* 2015 **468** 65-72
9. O'Shannessy DJ, Brigham-Burke M, Soneson KK, Hensley P, Brooks I. Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods. *Anal Biochem.* 1993 **212** 457-468

4

Role of glycine 221 in catalytic activity of Hyaluronan-Binding Protein 2

Fabian Stavenuiter[1], Eduard H.T.M. Ebberink[1], Koen Mertens[1,2] and Alexander B. Meijer[1,2]

J Biol Chem. 2017 292(15):6381-6388

[1]Department of Plasma Proteins, Sanquin Research, 1066 CX Amsterdam, The Netherlands
[2]Pharmaceutical Sciences, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands

ABSTRACT

Hyaluronan-Binding Protein 2 (HABP2) is a Ca^{2+}-dependent serine protease with putative roles in blood coagulation and fibrinolysis. A G221E substitution, known as the Marburg-I polymorphism, reportedly affects HABP2 function and has been associated with increased risk for cardiovascular disease. However, the importance of Gly-221 for HABP2 activity is unclear. Here, we used G221E, G221A, and G221S mutants to assess the role of Gly-221 in HABP2 catalysis. The G221E variant failed to activate single-chain urokinase-type plasminogen activator (scuPA), and the G221A and G221S variants displayed moderately reduced scuPA-activation. Activity toward the peptide substrate S-2288 was markedly decreased in all HABP2 variants, with G221E being the most defective and G221A the least defective. In the absence of Ca^{2+}, S-2288 cleavage by wild-type HABP2 was Na^+-dependent, with K_m decreasing from 3.0 to 0.6 mM upon titration from 0 to 0.3 M Na^+. In the presence of 5 mM Ca^{2+}, K_m was further reduced to 0.05 mM, but without appreciable contribution of Na^+. At physiological concentrations of Na^+ and Ca^{2+}, the three HABP2 variants, and particularly G221E, displayed a major K_m increase for S-2288. Chemical footprinting revealed that Ile-16 is significantly less protected from chemical modification in G221E than in wild-type HABP2, suggesting impaired insertion of the N-terminus into the G221E protease domain, with a concomitant impact on catalytic activity. Homology modeling suggested that the Glu-221 side chain could sterically hinder insertion of the N-terminus into the HABP2 protease domain, helping to explain the detrimental effects of Glu-221 substitution on HABP2 activity.

INTRODUCTION

Hyaluronan-Binding Protein 2 (HABP2) is a Ca^{2+}-dependent protease that belongs to the family of serine proteases. The single-chain zymogen is activated by a single cleavage, resulting in a light chain consisting of the catalytic domain with the typical chymotrypsin fold, and a heavy chain that comprises three EGF-like domains and a kringle domain.[1,2] HABP2 is also known as Factor Seven Activating Protease (FSAP). This, however, proved to be a misnomer, because coagulation factor VII is a remarkably poor substrate for FSAP/HABP2.[3] In contrast, HABP2 does effectively activate single-chain urokinase-type plasminogen activator (scuPA), suggesting a role in fibrinolysis.[3-5] HABP2 has further been reported to cleave tissue-factor pathway inhibitor (TFPI).[6] Interestingly, HABP2/FSAP knock-out mice display a mild bleeding phenotype, suggesting a role in hemostasis in relation to TFPI cleavage.[7]

As a serine protease with the chymotrypsin fold, the protease domain of HABP2 contains a dual antiparallel beta-barrel architecture.[8,9] In the beta-barrel cleft resides the catalytic triad comprising the residues Ser-195, His-57 and Asp-102 (chymotrypsin numbering is used throughout this paper). In extension of the protein core, there are eight unique surface loops. These loops appear more variable between serine proteases as they are involved in substrate recognition and allosteric regulation of the catalytic activity. For instance, the so-called 70-loop of HABP2 may interact with Ca^{2+} which is indispensable for allosteric regulation of the catalytic activity in coagulation factor IX and X.[2,10-12] In contrast, the 220-loop of thrombin, protein C, factor IX and factor X is involved in Na^{+}-binding.[13-16] The 220-loop is particularly important as it, together with the 180-loop, forms the primary S1 specificity pocket involved in recognition of the principle (P1) substrate residue (Arg in scuPA for HABP2). Moreover, the 220-loop is closely situated closely to the N-terminus insertion pocket.[8,9]

Whether or not the 220-loop has similar implications in HABP2 remains unclear. The HABP2 220-loop is considerably shorter than that of the typical Na^{+}-binding proteases (see Fig. 7A in the discussion section). On the other hand, a polymorphism in the HABP2 gene which encodes the amino acid substitution G221E in the 220-loop (G534E in HABP2 numbering) has been associated with an increased risk for cardiovascular disease or venous thrombosis.[17-19] This mutation has been designated as the Marburg-I polymorphism, and has been reported to cause reduced activity.[17] While the role of Gly-221 remains unclear, the above-mentioned findings indicate that G221E might not merely be a polymorphism and attributes a specific function to the 220-loop. In the present paper, we address the role of Gly-221 with particular reference to the regulation of HABP2 catalytic activity.

RESULTS

Construction of HABP2 derivatives

To elucidate the role of residue 221 for HABP2 function, we initially constructed three HABP2 variants in a WT-HABP2 background. In these variants, Gly-221 was replaced by either the small apolar alanine, the small polar serine, or by the negatively charged glutamic acid. Protein expression studies revealed that HABP2^{G221A} and HABP2^{G221S} were, like WT-HABP2, highly sensitive to autocatalytic inactivation (data not shown). Purified HABP2^{G221E} was, however, fully processed into a two-chain form (Fig. 1A). N-terminal sequencing identified Ile-16 as the N-terminus of the 25 kDa light chain, demonstrating that two-chain HABP2^{G221E} had been cleaved at the authentic activation site.

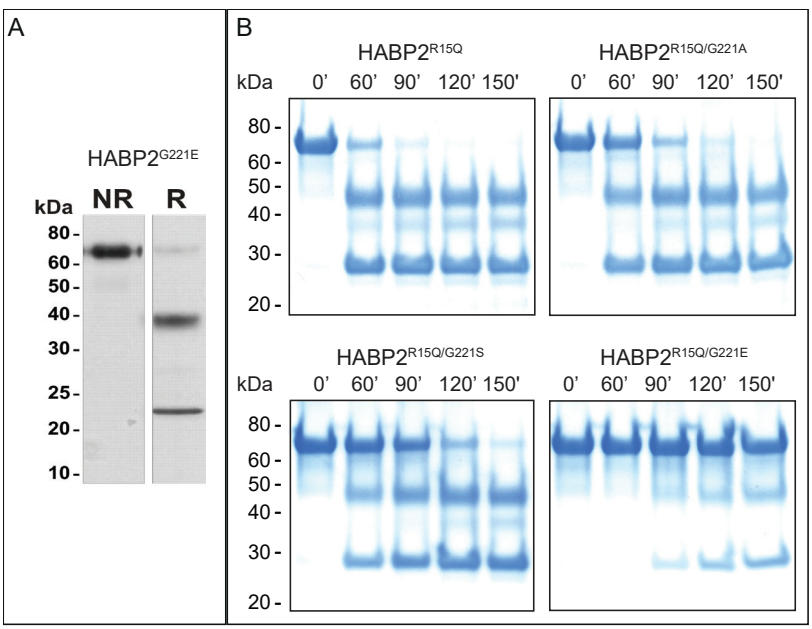


Figure 1: **SDS-PAGE analysis the HABP2 variants employed in this study. (A)** Purified HABP2^{G221E} (1 µg/lane) was analysed on 10% SDS-PAGE under reducing and non-reducing conditions. Protein bands were visualized via silver staining. **(B)** 0.7 µM HABP2^{R15Q}, HABP2$^{R15Q/G221A}$, HABP2$^{R15Q/G221S}$ and HABP2$^{R15Q/G221E}$ were incubated with 3 nM thermolysin for the indicated time intervals at 30 °C in 10 mM MES (pH 5.0), 150 mM NaCl, 5% (v/v) glycerol, 0.01% (v/v) Tween-80, 100 µM $CaCl_2$, and 50 nM $ZnCl_2$. Aliquots of the samples were separated on 4-12% SDS-PAGE and the protein bands were visualized by CBB staining. The position of the protein bands of the molecular weight marker is shown on the left.

We have previously demonstrated that the inactivation of HABP2 can be controlled by mutating the natural site of activation into a cleavage site for the protease thermolysin (HABP2^{R15Q}).[3] We therefore constructed thermolysin cleavage sites in all the HABP2 variants. After introduction of this mutation, the variants were purified as a single-chain molecule. There was, however, a differential effect of thermolysin on these HABP2 derivatives. Whereas HABP2^{R15Q} was fully activated by thermolysin within 90 minutes, HABP2$^{R15Q/G221A}$ and HABP2$^{R15Q/G221S}$ required a prolonged incubation time with thermolysin to obtain the two-chain molecule. An even longer incubation with thermolysin was required to proteolytically process HABP2$^{R15Q/G221E}$. However, this also led to additional cleavages within the light chain

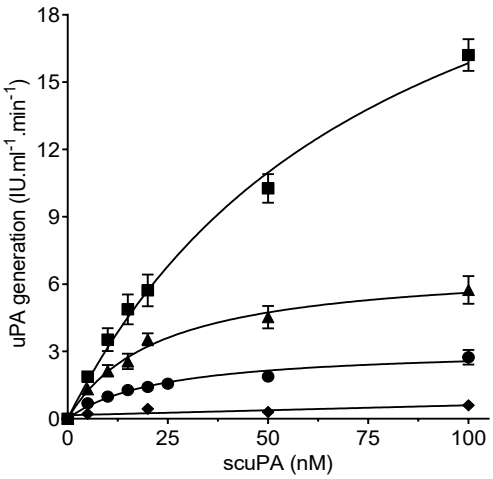


Figure 2: **Activation of scuPA by the HABP2 variants.** 0.25 nM of activated HABP2^{R15Q} (squares), HABP2$^{R15Q/G221A}$ (triangles), HABP2$^{R15Q/G221S}$ (circles), and HABP2^{G221E} (diamonds), were incubated with scuPA for 10 min at 37 °C in 100 mM NaCl, 0.1% (w/v) BSA, 0.01% (v/v) Tween-80, 5 mM $CaCl_2$, and 20 mM Tris-HCl (pH 7.5). Aprotinin was added to a final concentration of 125 nM to stop scuPA cleavage. The amount of generated uPA was measured employing S-2244 as described in Experimental Procedures. Data represent the mean ± S.D. of at least three experiments.

of this variant. We were therefore unable to obtain two-chain HABP2$^{R15Q/G221E}$ (Fig. 1B). These observations show that the nature of the amino acid at position 221 affects the activation of the HABP2 variants by thermolysin. To examine the role of amino acid substitution at 221, we employed in this study HABP2$^{R15Q/G221S}$, HABP2$^{R15Q/G221A}$, and HABP2^{R15Q} which can be specifically activated by thermolysin, as well as the constitutively active HABP2^{G221E}.

HABP2 activity towards macromolecular substrate scuPA

Each HABP2 variant was incubated with increasing concentrations of scuPA. The amount of uPA generated after 10 minutes was assessed employing the uPA-specific substrate S-2444 (Glu-Gly-Arg-*p*-nitroanilide). The data revealed that there is no detectable cleavage of scuPA in the presence of HABP2^{G221E} under these conditions

Table I. **Effects of Na^+ and Ca^{2+} on S-2288 hydrolysis.** The table below displays the kinetic constants of $HABP2^{R15Q}$ for hydrolysis of S-2288 either in presence or absence of Na^+ and Ca^{2+}. Constants were obtained from fitting the data of Fig. 3 using the Michaelis-Menten equation.

Na^+ (150 mM)	Ca^{2+} (5 mM)	K_m (mM)	$10^{-2} \times k_{cat}$ (min^{-1})	$10^{-2} \times (k_{cat} / K_m)$ ($min^{-1}\ mM^{-1}$)
-	-	3.13 ± 0.3	5.5 ± 0.3	1.8
+	-	0.64 ± 0.05	6.0 ± 0.1	9.4
-	+	0.06 ± 0.03	4.5 ± 0.3	75.0
+	+	0.05 ± 0.02	4.0 ± 0.2	79.2

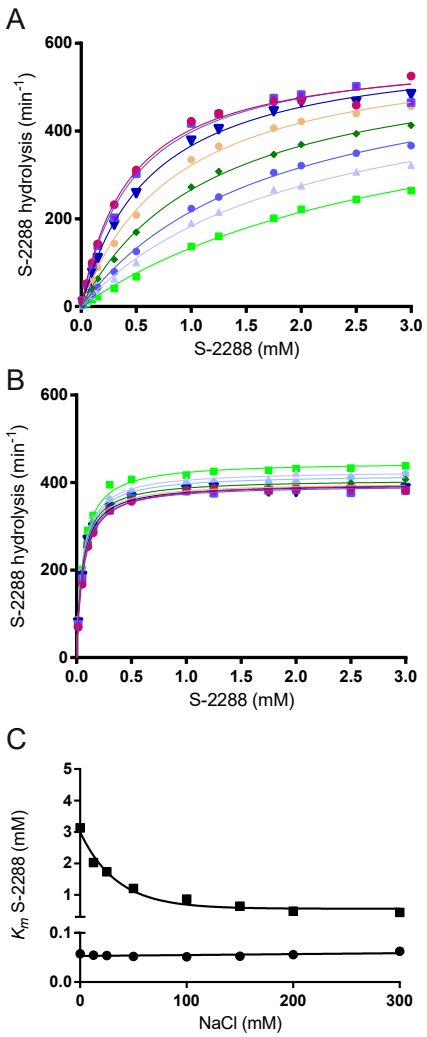


Figure 3: **Effects of Na^+ and Ca^{2+} on S-2288 hydrolysis by activated $HABP2^{R15Q}$. (A)** Displayed are the hydrolysis rates at varying concentrations of Na^+ in the absence of Ca^{2+}. S-2288 substrate conversion was assessed by incubation of 10 nM of activated $HABP2^{R15Q}$ at 37 °C in 50 mM Tris-HCl (pH 8.0), 5 mM EDTA and 0.1% (v/v) Tween-80. Na^+ concentrations are 0 mM (green squares), 12.5 mM (light blue triangles), 25 mM (blue circles), 50 mM (green diamonds), 100 mM (light orange circles), 150 mM (blue triangles), 200 mM (purple squares) and 300 mM (pink circles). Choline chloride was used to maintain the ionic strength at 1.5 M. Data represent mean values of three experiments; S.D. values were below 10% of the mean. **(B)** Using the same experimental set-up, S-2288 hydrolysis was assessed at varying concentrations of Na^+ in the presence of 5 mM Ca^{2+} instead of 5 mM EDTA. **(C)** Calculated K_m values of $HABP2^{R15Q}$ towards S-2288 either in absence (squares) or presence (circles) of 5 mM Ca^{2+}.

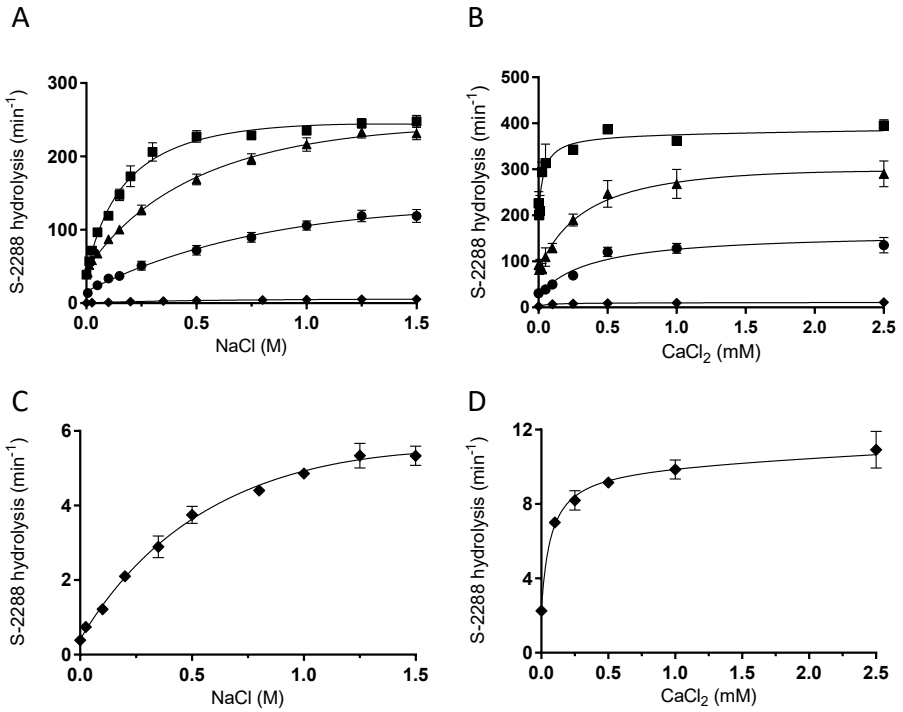


Figure 4: **Effects of Na^+ and Ca^{2+} on S-2288 hydrolysis by HABP2 variants. (A)** The effect of Na^+ on the activity of the HABP2 derivatives is shown in the absence of Ca^{2+}. 17.5 nM of activated $HABP2^{R15Q}$ (squares), $HABP2^{R15Q/G221A}$ (triangles), $HABP2^{R15Q/G221S}$ (circles) and 350 nM of $HABP2^{G221E}$ (diamonds) were incubated with varying concentrations of NaCl at 37 ºC in 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, and 0.1% (v/v) Tween-80. Choline chloride was added to keep the salt concentration constant at 1.5 M. S-2288 was added to a final concentration of 625 µM and its hydrolysis was assessed as described in Experimental Procedures. **(B)** The effect of Ca^{2+} on the activity of the HABP2 variants. 17.5 nM of activated $HABP2^{R15Q}$ (squares), $HABP2^{R15Q/G221A}$ (triangles), $HABP2^{R15Q/G221S}$ (circles) and 350 nM of $HABP2^{G221E}$ (diamonds) were incubated with varying concentrations of $CaCl_2$ in 150 mM NaCl, 50 mM Tris-HCl (pH 8.0), 0.1% (v/v) Tween-80. **(C, D)** Separate graphs are given to better display titration curves for $HABP2^{G221E}$.

(Fig. 2). $HABP2^{R15Q/G221A}$ and $HABP2^{R15Q/G221S}$ displayed slower scuPA activation than $HABP2^{R15Q}$. The data in Fig. 2 suggest that these mutants display a combination of lower K_m and k_{cat}, making the net effect on catalytic efficiency activity relatively small (see also Table II). Nevertheless, these data point to the possibility of changes in scuPA binding to HABP2 associated with mutations in position 221.

Effect of Na^+ on HABP2^{R15Q} activity in the presence and absence of Ca^{2+}

The catalytic activity of HABP2^{R15Q} towards the peptide substrate S-2288 (Ile-Pro-Arg-*p*-nitroanalide) was assessed at varying Na^+ concentrations in the presence or absence of Ca^{2+} (Fig. 3). In the absence of Ca^{2+}, the activity of HABP2^{R15Q} increased with increasing Na^+ concentrations (Fig. 3A). However, the effect of Na^+ was small in the presence of Ca^{2+} (Fig. 3B). Calculation of kinetic parameters showed that Na^+ causes a 5-fold drop in K_m in the absence of Ca^{2+}, while K_m proved independent of Na^+ when Ca^{2+} was present (Fig. 3C). As summarized in Table I, Na^+ and Ca^{2+} did predominantly affect K_m, but not k_{cat}. In terms of catalytic efficiency (k_{cat}/K_m) Na^+

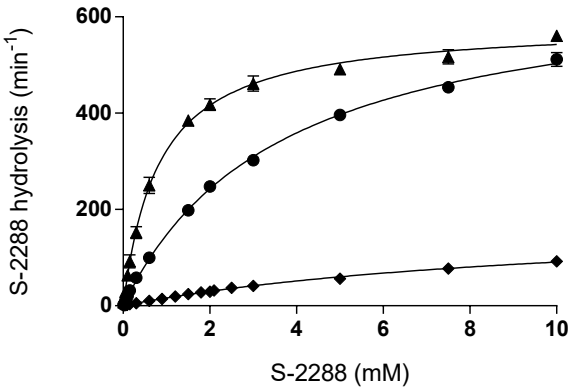


Figure 5: **Hydrolysis of S-2288 by the HABP2 derivatives.** 17.5 nM HABP2$^{R15Q/G221A}$ (triangles), 35 nM HABP2$^{R15Q/G221S}$ (circles) and 70 nM HABP2^{G221E} (diamonds) were incubated with different concentrations of S-2288 at 37 °C in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM $CaCl_2$ and 0.1% (v/v) Tween-80. The amount of cleaved S-2288 was measured as described in Experimental Procedures. Data represent the mean ± S.D. of at least three experiments.

Table II. **Effects of Gly-221 substitutions on scuPA activation and S-2288 hydrolysis.** Given are the calculated kinetic constants of the HABP2 variants for the proteolytic cleavage of scuPA and hydrolysis of S-2288. These constants were obtained by fitting data of respectively Fig. 2, Fig. 3 and Fig. 5 to the Michaelis-Menten equation. Due to the low activity of HABP2^{G221E} towards scuPA, kinetic constants thereof could not be determined (nd).

	hydrolysis of S-2288			activation of scuPA	
	K_m (mM)	$10^{-2} \times k_{cat}$ (min^{-1})	$10^{-2} \times (k_{cat} / K_m)$ (min^{-1} mM^{-1})	V_{max} (IU ml^{-1} min^{-1})	K_m (nM)
HABP2^{R15Q}	0.05 ± 0.02	4.0 ± 0.2	79.2	28 ± 3	80 ± 16
HABP2$^{R15Q/G221A}$	0.8 ± 0.03	5.9 ± 0.1	8.5	6.9 ± 0.6	23 ± 5
HABP2$^{R15Q/G221S}$	3.7 ± 0.1	6.8 ± 0.1	2.5	3.2 ± 0.3	24 ± 5
HABP2^{G221E}	10 ± 0.4	1.8 ± 0.1	0.2	nd	nd

stimulated HABP2 activity by 5-fold, while the effect of Ca^{2+} was 8-fold higher (Table I). Thus, in comparison with the magnitude of the effect of Ca^{2+}, the effect of Na^{+} seems limited. Control experiments (data not shown) indicated that in the absence of Ca^{2+}, Li^{+} had no effect on the amidolytic activity of HABP2. The same holds for choline chloride, which was used to keep the ionic strength constant in the Na^{+}-titration studies.

Amidolytic activity of the HABP2 variants

The activity HABP2 variants was examined by Na^{+}- and Ca^{2+}-titration experiments. The activity of the mutants was dependent on the amino acid in position 221, in the order Gly>Ala>Ser>Glu (Fig. 4A and B). Like $HABP2^{R15Q}$, increasing Na^{+}-

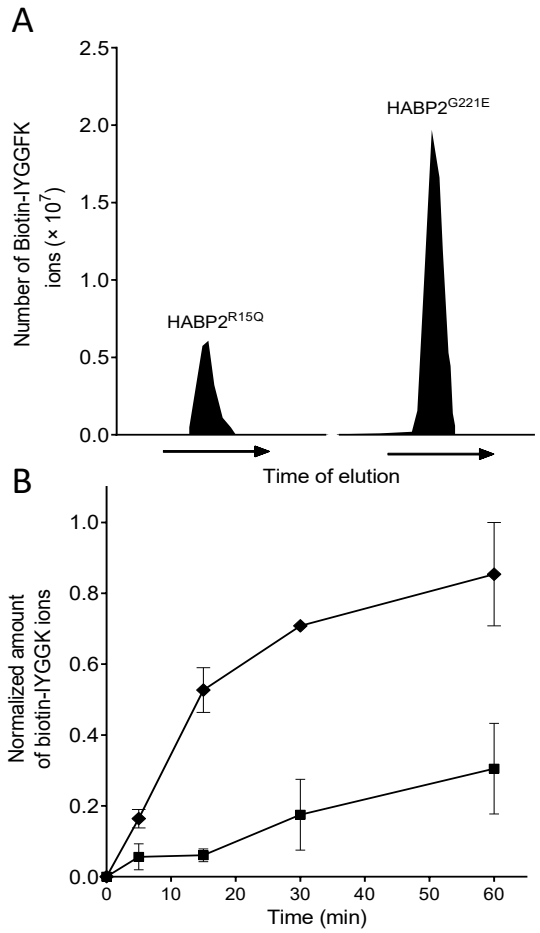


Figure 6: **Reduced protection from chemical modification of the amino group of Ile-16 in $HABP2^{G221E}$.** $HABP2^{R15Q}$ and $HABP2^{G221E}$ were incubated with sulpho-NHS-LC-biotin for different periods of time. Peptides of the HABP2 derivatives were obtained and separated on a C18 liquid chromatography column which was connected to a mass spectrometer as described in materials and methods. **(A)** Shown are the reconstructed ion chromatograms of the eluted peptide representing the modified N-terminus of $HABP2^{R15Q}$ and of $HABP2^{G221E}$ after 15 minutes of incubation with sulpho-NHS-LC-biotin. **(B)** Shown are the total number of modified N-terminal peptide ions of $HABP2^{R15Q}$ (squares) and $HABP2^{G221E}$ (diamonds) normalized to the maximum number of modified ions that were obtained for $HABP2^{G221E}$ after 60 minutes of incubation with sulpho-NHS-LC-biotin (see materials and methods). The data show the average of two independent experiments. The error bars show the deviation between the two experiments.

concentrations enhanced the activity of the HABP2 variants (Fig. 4A, C). Similarly, all variants displayed similar response to Ca^{2+} in the presence of Na^{+} (Fig. 4B, D), the optimal rate of substrate hydrolysis being reached at approximately 1 mM for all variants.

The kinetic parameters for S-2288 hydrolysis were then assessed in the presence of 5 mM $CaCl_2$ and 150 mM NaCl (Fig. 5, Table II). These data show that replacement of Gly in position 221 greatly reduces the catalytic efficiency. The most affected variant was $HABP2^{R15Q/G221E}$, which proved 400-fold less efficient than $HABP2^{R15Q}$. Because substitution of Gly-221 predominantly affects K_m, these data are compatible with perturbation of the interaction between active site of HABP2 and the substrate S-2288. Thus Gly-221 in the 220-loop is beneficial for HABP2 activity towards both the peptide substrate S-2288 and the macromolecular substrate scuPA (Table II).

Insertion of Ile-16 in the protease domain of $HABP2^{G221E}$

Following proteolytic activation of $HABP2^{R15Q}$ and $HABP2^{G221E}$, insertion of the newly formed N-terminus (Ile-16) was probed by incubating with sulpho-NHS-LC-biotin that specifically reacts with free, solvent-exposed amino groups.[20,21] The biotinylated proteins were proteolytically processed with trypsin and analysed by mass spectrometry. After 5 minutes of incubation with sulpho-NHS-LC-biotin, a biotinylated N-terminal peptide of HABP2 was identified (i.e. Biotin-IYGGFK, m/z 512.271) for both activated $HABP2^{R15Q}$ and activated $HABP2^{G221E}$. The area under the curve of the reconstructed ion chromatograms shows that the degree of biotinylation at Ile-16 was at each time point higher for $HABP2^{G221E}$ than for $HABP2^{R15Q}$ (Fig. 6).

DISCUSSION

The aim of the present study was to address the role of Gly-534 in the serine protease HABP2, and in particular the functional implication of the G534E substitution. This naturally occurring variant is encoded by the Marburg-I polymorphism in the HABP2 gene, and has been associated with increased cardiovascular risk. In chymotrypsin numbering, the Marburg-I polymorphism affects position 221 and thus is part of the characteristic 220-loop that participates in allosteric regulation of the catalytic activity in a variety of serine proteases.[13-16,22] This raised the possibility that this loop, and in particular residue 221 therein, would serve a similar role in HABP2.

In related serine proteases such as activated protein C, factor IXa and factor Xa, allosteric regulation involves interplay between the S1-site in the specificity pocket, the Ca^{2+}-binding site in the protease domain, and Na^{+}-binding that involves

the 220-loop.[14-16] In HABP2, the 220-loop differs from that in other mammalian serine proteases, in that it is two amino acid residues shorter (see Fig. 7A). Moreover, HABP2 has Pro in position 225, which is typical for trypsin-like proteases, whereas most Na^+-binding serine proteases have Tyr in this position.[22] In view of these differences, it is surprising that the amidolytic activity of activated HABP2 is sensitive not only to Ca^{2+}, but also to Na^+ (Figs. 3 and 4). The catalytic efficiency (k_{cat}/K_m) was approximately 5-fold increased by for Na^+ alone, and 42-fold by Ca^{2+} alone (see Table I). In the presence of Ca^{2+}, the contribution of Na^+ proved virtually absent, presumably because it is small in comparison to the Ca^{2+}-response (Fig. 3B, Table I). In this regard, HABP2 differs from for instance factor IXa, wherein Na^+ and Ca^{2+} contribute to a similar extent,[14] or activated protein C, wherein the Na^+-response predominates.[15] The limited response to Na^+ in the presence of Ca^{2+} makes it difficult to conclude whether or not HABP2, like activated protein C, factor IXa and factor Xa,[14-16] displays any thermodynamic linkage between the putative Na^+- and the Ca^{2+}-binding sites. It is evident, however, that Na^+ induces a 5-fold reduction of K_m for the peptide substrate S-2288 (Fig. 3C). Because k_{cat} remains essentially unchanged, the drop in K_m suggests that Na^+ substantially facilitates the interaction of HABP2 with the peptide substrate. This effect seems negligible, however, in the presence of Ca^{2+}, which reduces K_m by a further 12-fold (Fig. 3B). The Na^+-sensitivity of HABP2 implies that it is a Na^+-binding protease, in a manner that is highly reminiscent to related serine proteases that display Na^+-binding to the 220-loop. Because the effect of Na^+ is no longer apparent at physiological Ca^{2+} concentrations (Fig. 3C, Table I), this is unlikely to be of physiological significance. In this respect HABP2 may be classified as being Na^+-*binding*, but not Na^+-*dependent* in the strict sense.

As for the role of Gly in position 221, it is evident that the substitution into Ala or Ser leads to a moderate reduction of proteolytic activity of HABP2 towards scuPA, while mutation into Glu, the Marburg-I variant, displays by far the most prominent defect (Fig. 2, Table II). Similar conclusions can be drawn with respect to amidolytic activity towards S-2288 (Fig. 5), although in terms of catalytic efficiency the difference between mutants seems more pronounced (Table II). The G221A and G221S substitutions predominantly increased K_m, which seems indicative for perturbed substrate interaction, while the effect of the G221E variant seems more complex, affecting both K_m and k_{cat}. We considered the possibility that the substitutions may severely affect the response to Na^+ and Ca^{2+}, but titration experiments did not provide any appreciable evidence in this direction (Fig. 4). Thus, the severely impaired activity of $HABP2^{G221E}$ is unlikely to be due to a major disruption in the interaction between the S1-specificity pocket and the binding sites for Na^+ and/or Ca^{2+}. Another interaction with the S1-pocket in serine proteases occurs upon proteolytic activation, via an electrostatic interaction between Asp-194 and the nascent new N-terminus Ile-16.[9]

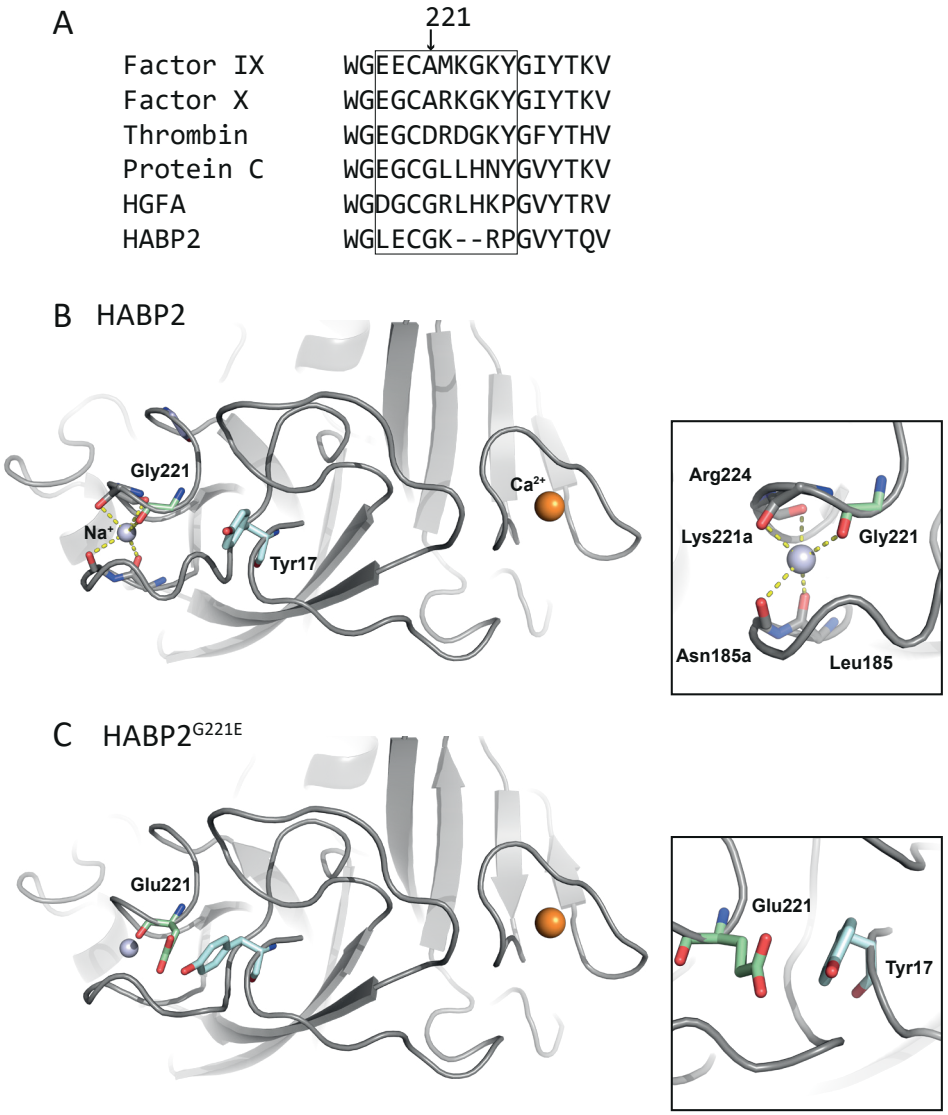


Figure 7 *(figure legend on next page)*

The observation that in HABP2^{G221E} the N-terminus is significantly more exposed than in normal HABP2 (Fig. 6) suggests that HABP2^{G221E} lacks the stabilization of the S1-pocket that normally results from N-terminus insertion. As such, the activated form of HABP2^{G221E} may predominantly be inactive, presumably due to a shift of the equilibrium between closed (E*) and open (E) forms,[23] or in the continuum between zymogen- and proteinase-like states.[24]

In absence of a crystal structure of HABP2, it remains difficult to fully

Figure 7: **Position 221 in the HABP2 protease domain. (A)** The 220-loop (delineated by a box) of HABP2 is shorter than that of other serine proteases as visualized by the primary sequence alignment. Except for Cys-220, the 220-loop residues of HABP2 are poorly conserved with those of factor IX, X, thrombin and protein C. **(B)** Models of the HABP2 protease domain were constructed with Modeller 9v7 employing the crystal structure of HGFA as a template (PDB code: 1YC0).[26] The presence of the Na^+-ion was modelled according to its position in the FXa crystal structure (PDB code: 2BOK).[32] A representative image is shown that also comprises the Ca^{2+}-binding site, 220-loop, S1-pocket and N-terminus insertion site. The modelled Na^+-ion is shown as a light blue sphere, the Ca^{2+}-ion is shown as an orange sphere. A close up is given of the 180-loop and 220-loop in interaction with a Na^+-ion. Putative coordination of the Na^+-ion by amino acid carbonyl groups of Leu-185, Asn185a, Gly-221, Lys-221a and Arg-224 is indicated by dashed lines. Tyr-17 at the N-terminus of the protease domain is indicated in cyan, Gly-221 is shown in a green colour stick representation. **(C)** The model of $HABP2^{G221E}$ displays the potential steric and electrostatic repulsion between Glu-221 and Tyr-17 at the N-terminus insertion site. On the right, a close-up is shown of Glu-221 and Tyr-17 in respectively green- and cyan-coloured stick representations. All images were generated using PyMol (Schrödinger, LLC).

appreciate the role of Gly-221 and the effect of the mutations described in the present paper. Of all human serine proteases, hepatocyte growth factor activator (HGFA) exhibits the highest primary sequence homology with full-length HABP2.[1] To assess the structural role Gly-221, we constructed a model of the catalytic domain of HABP2 based on that of HGFA by employing comparative homology modeling.[25,26] Fig. 7B shows the structure of the active site region of HABP2, with the putative position of the Ca^{2+} and Na^+ ions. As for Na^+-binding, five out of six coordinating ligands could be provided by the carbonyl O-atoms of residues Gly-221, Lys-221a, Arg-224 from the 220-loop, and Leu-185 and Asn-185a from the opposite 180-loop (Fig. 7B). While in other Na^+-binding serine proteases the coordination for Na^+-binding is disrupted by Pro in position 225,[22] HABP2 may be an exception because its 220-loop is shorter, and thereby less sensitive to destabilization by the side chain at position 225. The model suggests that in $HABP2^{G221E}$ the putative Na^+-coordination may be conserved (Fig. 7C), which seems compatible with the observed Na^+-response of this variant (Fig. 4C). The model further suggests that Glu in position 221 could introduce a steric clash with Tyr-17, thus impeding the insertion of the new N-terminus Ile-16 after HABP2 activation (Fig. 7C), and the concomitant conversion of the zymogen-like form into the proteinase-like conformation.

The findings of this study show that Gly-221 is of major importance for the catalytic activity of HABP2. As such, the significance of position 221 is shared with other serine proteases like protein C, factor IX and thrombin. For thrombin, mutation of the Asp-221 into a glutamic acid, which is highly similar in properties to the former amino acid residue, already leads to a severe defect in fibrinogen clotting.[27] Similarly, the G221R mutation in protein C leads to an impaired biological activity of

this protease.[28] In contrast, alanine to aspartic acid exchange in coagulation factor IX does not affect its clotting activity or its binding to macromolecular substrates at all.[29] However, mutation of the same alanine into a valine severely impairs activity, which has been attributed to obstruction of N-terminus insertion by the branched side chain.[29] Our study proposes a similar role of position 221 in the 220-loop on N-terminus insertion in HABP2.

In view of the numerous substrates that have been proposed for HABP2[3-7,10] it remains difficult to predict the phenotype associated with Gly-221 substitution. It remains plausible that HABP2 is a promiscuous enzyme that activates more substrates than scuPA alone, and thus may have more roles than stimulation of fibrinolysis. However, irrespective of the physiological function of HABP2, it is evident that the G221E substitution as encoded by the Marburg-I polymorphism results in a severely dysfunctional protein.

EXPERIMENTAL PROCEDURES

Chemicals

All chemical were from Sigma-Aldrich (Zwijndrecht, The Netherlands) unless otherwise stated. Sulpho-NHS-LC-biotin was from QB Perbio (Tattenhall, Cheshire, UK). Acetonitrile and ultra-pure water, which were employed for mass spectrometry, were obtained from Biosolve B.V. (Valkenswaard, The Netherlands).

Recombinant HABP2 derivatives and their activation by thermolysin

Recombinant $HABP2^{R15Q}$ is described in Stavenuiter *et al*.[3] The thermolysin HABP2 variants carrying amino acid substitution at position 221 were constructed employing QuikChange mutagenesis (Stratagen, La Jolla, CA) using $HABP2^{R15Q}$ DNA as a template and the appropriate primers. Recombinant proteins were produced, purified, activated and quantified as described.[3] The HABP2 variants were activated by incubating 0.7 µM of the protein with 3 nM thermolysin for 2 hours at 30 ºC in buffer containing 10 mM MES (pH 5.0), 150 mM NaCl, 5% (v/v) glycerol, 0.01% (v/v) Tween-80, 100 µM $CaCl_2$, and 50 nM $ZnCl_2$. Thermolysin was inhibited by the addition of phosphoramidon disodium salt to a final concentration of 10 µM. Because of its inherent instability, HABP2 was freshly activated for all individual experiments. SDS/PAGE was performed to check whether activation was complete (>90-95%). Molar concentrations of activated HABP2 were based on protein concentrations prior to activation.

Hydrolysis of S-2288 by HABP2 derivatives and the determination of uPA concentration

The initial rates of *p*-nitroaniline release from S-2288 were measured by following the absorbance at 405 nm using a Rosys-Anthos Lucy 3 photometer (Anthos Labtec Instruments GmbH, Wals, Austria). To calculate the concentration of *p*-nitroaniline an extinction coefficient of 9.9 mM^{-1} cm^{-1} was used. The concentration of uPA, obtained from scuPA cleavage by the HABP2 derivatives, was assessed by adding S-2444 to the reaction mixture (final concentration 375 µM).[3] The initial rates of *p*-nitroaniline release were assessed by monitoring the absorbance at 405 nm in time as described above. The results were corrected for the background levels of *p*-nitroaniline in the absence of the HABP2 derivatives. Absorbance changes were converted into increase in *p*-nitroaniline concentration and divided by the enzyme concentration, to yield S-2288 hydrolysis rates in min^{-1}. The concentration of uPA was determined using a calibration curve of the HMW uPA international standard 87/594 (NIBSC, South Mimms, England).

Determination of steady states kinetic constants and inhibition kinetics

Kinetic constants of S-2288 hydrolysis and scuPA activation by the HABP2 derivatives were assessed by assuming Michaelis-Menten-like behavior for the enzymatic reactions.[30] Graphpad software was employed to fit the data to the Michaelis-Menten equation employing non-linear regression.

Biotinylation of the free amino group of Ile-16

350 nM of activated HABP2^{R15Q} or HABP2^{G221E} was incubated with 25 mM sulpho-NHS-biotin in 100 mM NaCl, 1 mM $CaCl_2$, 0.02% (v/v) Tween-80, 50 mM HEPES (pH 7.5). At different time points the reactions were stopped by the addition of lysine (125 mM). Proteins were subsequently separated on precast NuPage 4-12% Bis-Tris gels (Invitrogen, Breda, The Netherlands) under reducing conditions. The protein bands were subsequently visualized employing Coomassie Brilliant Blue (CBB) staining. Bands of interest were excised and processed for in-gel digestion according to the method of Shevchenko.[31] Briefly, bands were washed with a buffer containing 50 mM ammonium bicarbonate (pH 7.9) followed by a wash with 50% (v/v) acetonitrile. This step was repeated three-times. Cysteine bonds were subsequently reduced with 10 mM dithiothreitol for 1 h at 56 °C, and alkylated with 50 mM iodoacetamide for 45 min at RT in the dark. After two subsequent wash/dehydration cycles the bands were dried for 10 min in a vacuum centrifuge (Thermo Fisher Scientific Inc., Bremen, Germany) and incubated overnight with 0.06 µg/µl trypsin at 37 °C. Peptides were eluted from the gel with 1% (v/v) formic acid and twice with 50% (v/v) acetonitrile, 5% (v/v) formic acid. The three elution fractions were pooled and concentrated to 20 µl in a vacuum centrifuge.

Mass spectrometry and data analysis

The above-described peptides were separated using a reversed-phase C18 column (50 µm x 40 cm, 5 µm particles) (Nanoseparations, Nieuwkoop, The Netherlands) running at 100 nl/min with a 1 hour gradient from 0% to 35 % (v/v) acetonitrile with 0.1% (v/v) acetic acid. The peptides were sprayed directly from the column into the LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific Inc., Bremen, Germany) using a nanoelectrospray source with a spray voltage of 1.9 kV. The LTQ was operated in a data-dependent mode by performing collision induced dissociation in the ion-trap (35% normalized collision energy) for the five most intense precursor ions selected from each full scan in the Orbitrap (350-2000 m/z, resolving power 30.000). An isolation width of 2 Da was used for the selected ions (charge ≥ 2) and an activation time of 30 ms. Dynamic exclusion was activated for the MS/MS scan with

a repeat count of 1 and exclusion duration of 30 s. Peptides were identified employing a Sequest search against the human entries in the NCBI database utilizing Proteome Discoverer 1.0 software (Thermo Scientific, Bremen, Germany). The same software was used to obtain the peak area of the reconstructed ion chromatograms of the peptide precursor ions for quantification purposes. Direct comparison of the tryptic digests obtained from gel is not possible due to a differential loss of the peptides during the peptide purification procedure. Therefore, the areas of the reconstructed ion chromatograms of control peptides present in each sample were used as a reference to correct for peptide loss.

Author contributions

FS constructed and purified the HABP2 variants, FS performed and analyzed experiments, EE constructed the HAPB2 models. FS, EE, KM and ABM wrote the manuscript. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

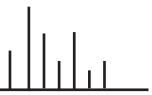
1. Choi-Miura NH, Tobe T, Sumiya J, Nakano Y, Sano Y, Mazda T, Tomita M. Purification and characterization of a novel hyaluronan-binding protein (PHBP) from human plasma: it has three EGF, a kringle and a serine protease domain, similar to hepatocyte growth factor activator. *J Biochem.* 1996 **119** 1157-1165.
2. Hunfeld A, Etscheid M, König H, Seitz R, and Dodt J. Detection of a novel plasma serine protease during purification of vitamin K-dependent coagulation factors. *FEBS Lett.* 1999 **456** 290-294.
3. Stavenuiter F, Dienava-Verdoold I, Boon-Spijker MG, Meijer AB, Mertens K. Factor Seven Activating Protease (FSAP): does it activate Factor VII? *J Thromb Haemost.* 2012 **10** 859-866.
4. Choi-Miura NH, Yoda M, Saito K, Takahashi K, Tomita M. Identification of the substrates for plasma hyaluronan binding protein. *Biol Pharm Bull.* 2001 **24** 140-143.
5. Römisch J, Vermöhlen S, Feussner A, Stöhr H. The FVII activating protease cleaves single-chain plasminogen activators. *Haemostasis* 1999 **29** 292-299.
6. Kanse SM, Declerck PJ, Ruf W, Broze G, Etscheid M. Factor VII-Activating Protease Promotes the Proteolysis and Inhibition of Tissue Factor Pathway Inhibitor. *Arterioscler Thromb Vasc Biol.* 2012 **32** 427–433
7. Subramaniam S, Thielmann I, Morowski M, Pragst I, Sandset PM, Nieswandt B, Etscheid M, Kanse SM. Defective thrombus formation in mice lacking endogenous factor VII activating protease (FSAP). *Thromb Haemost.* 2015 **113** 870-880
8. Di Cera, E. Serine proteases. *IUBMB Life.* 2009 **61** 510-515.
9. Page MJ, Di Cera E. Serine peptidases: classification, structure and function. *Cell Mol Life Sci.* 2008 **65** 1220-1236.
10. Kannemeier C, Feussner A, Stöhr HA, Weisse J, Preissner KT, Römisch J. Factor VII and single-chain plasminogen activator-activating protease: activation and autoactivation of the proenzyme. *Eur J Biochem.* 2001 **268** 3789-3796.
11. Mathur A, Zhong D, Sabharwal AK, Smith KJ, Bajaj SP. Interaction of factor IXa with factor VIIIa. Effects of protease domain Ca^{2+} binding site, proteolysis in the autolysis loop, phospholipid, and factor X. *J Biol Chem.* 1997 **272** 23418-23426.
12. Sabharwal AK, Padmanabhan K, Tulinsky A, Mathur A, Gorka J, Bajaj SP. Interaction of calcium with native and decarboxylated human factor X. Effect of proteolysis in the autolysis loop on catalytic efficiency and factor Va binding. *J Biol Chem.* 1997 **272** 22037-22045.
13. Pineda AO, Carrell CJ, Bush LA, Prasad S, Caccia S, Chen ZW, Mathews FS, Di Cera E. Molecular dissection of Na^{+} binding to thrombin. *J Biol Chem.* 2004 **279** 31842-31853.

14. Schmidt AE, Stewart JE, Mathur A, Krishnaswamy S, Bajaj SP. Na^+ site in blood coagulation factor IXa: effect on catalysis and factor VIIIa binding. *J Mol Biol.* 2005 **350** 78-91.
15. Schmidt AE, Padmanabhan K, Underwood MC, Bode W, Mather T, Bajaj SP. Thermodynamic linkage between the S1 site, the Na^+ site, and the Ca^{2+} site in the protease domain of human activated protein C (APC): Sodium ion in the APC crystal structure is coordinated to four carbonyl groups from two separate loops. *J Biol Chem.* 2002 **277** 28987-28995.
16. Underwood MC, Zhong D, Mathur A, Heyduk T, Bajaj SP. Thermodynamic linkage between the S1 site, the Na^+ site, and the Ca^{2+} site in the protease domain of human coagulation factor Xa: Studies on catalytic efficiency and inhibitor binding. *J Biol Chem.* 1996 **275** 36876-36884.
17. Etscheid M, Muhl L, Pons D, Jukema JW, König H, Kanse SM. The Marburg I polymorphism of factor VII activating protease is associated with low proteolytic and low pro-coagulant activity. *Thromb Res.* 2012 **130** 935-941.
18. Ireland H, Miller GJ, Webb KE, Cooper JA, Humphries SE. The factor VII activating protease G511E (Marburg) variant and cardiovascular risk. *Thromb Haemost.* 2004 **92** 986-992.
19. Weisbach V, Ruppel R, Eckstein R. The Marburg I polymorphism of factor VII-activating protease and the risk of venous thromboembolism. *Thromb Haemost.* 2007 **97** 870-872.
20. Gabant G, Augier J, Armengaud J. Assessment of solvent residues accessibility using three Sulfo-NHS-biotin reagents in parallel: application to footprint changes of a methyltransferase upon binding its substrate. *J Mass Spectrom.* 2008 **43** 360-370.
21. Azim-Zadeh O, Hillebrecht A, Linne U, Marahiel MA, Klebe G, Lingelbach K. Nyalwidhe J. Use of biotin derivatives to probe conformational changes in proteins. *J Biol Chem.* 2007 **282** 21609-21617.
22. Dang QD, Di Cera E. Residue 225 determines the Na^+-induced allosteric regulation of catalytic activity in serine proteases. *Proc Natl Acad Sci USA*. 1996 **93** 10653-10656.
23. Vogt AD, Chakraborty P, Di Cera E. Kinetic dissection of the pre-existing conformational equilibrium in the trypsin fold. *J Biol Chem.* 2015 **290** 22435-22445.
24. Kamath P, Huntington JA, Krishnaswamy S. Ligand binding shuttles thrombin along a continuum of zymogen- and proteinase-like states. *J Biol Chem.* 2010 **285** 28651-28658.
25. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993 **234** 779-815.
26. Shia S, Stamos J, Kirchhofer D, Fan B, Wu J, Corpuz RT, Santell L, Lazarus RA,

Eigenbrot C. Conformational lability in serine protease active sites: structures of hepatocyte growth factor activator (HGFA) alone and with the inhibitory domain from HGFA inhibitor-1B. *J Mol Biol.* 2005 **346** 1335-1349

27. Rouy S, Vidaud D, Alessandri JL, Dautzenberg MD, Venisse L, Guillin MC, Bezeaud A. Prothrombin Saint-Denis: a natural variant with a point mutation resulting in Asp to Glu substitution at position 552 in prothrombin. *Br J Haematol.* 2006 **132** 770-773.
28. Miyata T, Zheng YZ, Sakata T, Tsushima N, Kato H. Three missense mutations in the protein C heavy chain causing type I and type II protein C deficiency. *Thromb Haemost.* 1994 **71** 32-37.
29. Hamaguchi N, Roberts H, Stafford DW. Mutations in the catalytic domain of factor IX that are related to the subclass hemophilia Bm. *Biochemistry.* 1993 **32** 6324-6329.
30. Segel IH. Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems. 1975 John Wiley & Sons, Inc, New York.
31. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc.* 2006 **1** 2856-2860.
32. Schärer K, Morgenthaler M, Paulini R, Obst-Sander U, Banner DW, Schlatter D, Benz J, Stihle M, Diederich F. Quantification of cation-pi interactions in protein-ligand complexes: crystal-structure analysis of Factor Xa bound to a quaternary ammonium ion ligand. *Angew Chem Int Ed Engl.* 2005 **44** 4400-4404.

5



Probing activation-driven changes in coagulation factor X, IX and prothrombin by primary amine labelling

Eduard H.T.M. Ebberink[1], Nadia Freato[1], Floris P.J. van Alphen[1], Mariëtte G. Boon-Spijker[1], Maartje van den Biggelaar[1,2], Alexander B. Meijer[1,2] and Koen Mertens[1,2]

Manuscript in preparation

[1]Department of Plasma Proteins, Sanquin Research, 1066 CX Amsterdam, The Netherlands
[2]Pharmaceutical Sciences, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands

ABSTRACT

Factor IX (FIX), factor X (FX) and prothrombin are the zymogen forms of their activated counterparts FIXa, FXa and thrombin. Despite structural homology, thrombin, FXa and FIXa display variable proteolytic activity, with FIXa having remarkable low intrinsic activity. As in other chymotrypsin-like serine proteases, activation occurs by limited proteolysis. This endows the protease domain with a novel N-terminus that is thought to insert into the catalytic site to drive protease activity. The structural details of zymogen-to-protease conversion have remained poorly understood. Here we probe these structural rearrangements by primary amine labelling and mass spectrometry. Upon activation of prothrombin, several changes were observed in the surface loops of the catalytic domain, in particular in the so-called 220-loop. Lysine residues in this loop displayed a 4-fold reduced reactivity, suggesting a reorientation of these lysines due to activation. Similar changes were observed in the 220-loop of FX, but not in FIX. The same approach was used to assess the extent of labelling of the new N-terminus. In all three enzymes, the N-terminus became more protected upon incorporation of an irreversible inhibitor in the active site. In FIXa, the binding of this substrate analogue was accompanied by a decrease in reactivity of the 220-loop. This did not occur in FXa and thrombin. These data indicate that the zymogen-to-protease transition in FIXa differs from that of FXa and thrombin and this may involve dynamics of the 220-loop. In comparison with FXa and thrombin, FIXa seems more zymogen-like, which might explain its low intrinsic catalytic activity.

INTRODUCTION

Blood coagulation involves a self-amplifying cascade of sequential conversions of inactive zymogens into active serine proteases. The main proenzymes herein are factor VII (FVII), factor IX (FIX), factor X (FX) and prothrombin, which are converted by limited proteolysis into their active counterparts called FVIIa, FIXa, FXa, and thrombin, respectively.[1,2] These enzymes belong to the chymotrypsin superfamily of serine proteases.[3-5] Their catalytic domains display a dual antiparallel β-barrel architecture, in which the interface between the β-barrel domains encloses the catalytic triad, the oxyanion hole, and the substrate recognition pockets.[3,4] Apart from these common structure elements, the β-barrels carry eight surface loops that are variable between individual serine proteases. In most of the coagulation enzymes, these loops are longer than in chymotrypsin. Therefore they may be viewed as insertions that protrude from the protein core, in positions that facilitate participation in unique, protease-specific interactions, or in more general allosteric events. The so-called 60-loop and 140-loop (chymotrypsin numbering is used throughout this paper) directly flank the substrate binding cleft and thereby affect the interaction with substrates.[6] An allosteric role has been established for the 70-loop, which provides a Ca^{2+}-binding site in FVIIa, FIXa and FXa.[7-9] Similarly, the 220-loop, together with the 180-loop, comprises a Na^{+}-binding site that contributes to the catalytic activity of thrombin, FXa and, to a lesser extent, of FIXa.[10-12]

Limited proteolysis is an essential step in the conversion of chymotrypsin-like zymogens into active serine proteases. This involves a cleavage in the N-terminal region of the protease domain, between residue 15 and 16, resulting in a novel N-terminus, usually an isoleucine residue. The nascent N-terminus then inserts into a binding pocket by forming an ion-pair between Ile16 and Asp194. This drives the proper organization of the oxyanion hole and the substrate-binding site and of the Arg-binding S1-pocket that is adjacent to Ser195 in the catalytic triad.[3,4] In chymotrypsin, limited proteolysis alone seems sufficient to drive the protease domain to an active state with its active site 'open' and accessible to the substrate. For the proteases of the coagulation system, the transition of zymogen to active protease is more complex. In particular, thrombin has been extensively studied. Structural information and rapid kinetics have established that thrombin, despite being fully activated in terms of proteolytic processing, can adopt both protease-like and zymogen-like forms, which are in a dynamic equilibrium.[13-15] FVIIa is known to deviate from the classical view in that its N-terminus is accessible to chemical modification, and thus is not, or only partially, inserted into the active site.[16] Moreover, recent studies using hydrogen/deuterium exchange have demonstrated that FVII and FVIIa share the same conformation and that the transition into a protease-like conformation requires

the presence of an irreversible inhibitor in the S1-pocket in the active site.[17] This raises the question as whether coagulation factors downstream of FVII in the coagulation cascade display any proteolysis-driven conformational changes. In the present study, we addressed this question for FIX, FX and prothrombin.

Proteolytic activation was monitored using a combination of chemical footprinting and mass spectrometry. For this purpose, we employed a tandem-mass-tags (TMTs) approach, which enables a pairwise comparison between samples during zymogen activation. The TMT probes target primary amines, and therefore may provide information about changes in accessibility of lysine side chains and the N-terminus of the protease domain. We used the TMT approach previously to identify lysine residues that are involved in the function of coagulation factor VIII (FVIII),[18,19] von Willebrand Factor (VWF)[20] and receptor-associated protein (RAP).[21] We now used the same technique for an unbiased, time-resolved analysis of changes that occur upon the generation of the new N-terminus. The results suggest that FIXa is predominantly zymogen-like, whereas FXa and thrombin display a more enzyme-like pattern, in particular with regard to reactivity of lysine residues in their respective 220-loops.

RESULTS

Time-resolved lysine labelling of prothrombin and FX during activation

Prothrombin was activated by incubation with *Echis Carinatus* snake venom (ECV). Non-reduced SDS/PAGE (Fig. 1A) shows a pattern that is compatible with cleavage occurring at three positions: Arg155$^{(FII)}$, followed by Arg271$^{(FII)}$ and Arg320$^{(FII)}$.[14,22] After 1 hour, the predominant products were prethrombin-1 and fragment 1. Enzymatically active forms started to accumulated after 2 hours, and after 4 hours the mixture consisted of fragment-1, fragment-2, and the mature enzyme thrombin (A+B in Fig. 1A). In chymotrypsin numbering, Arg320$^{(FII)}$ corresponds with Arg15, thus representing the cleavage that generates the new N-terminus of the protease domain.

After incubation with ECV, lysine labelling was performed in order to probe changes in protein structure. Figure 1B shows the lysine reactivity of the prothrombin components in the reaction mixture in our TMT approach. Prothrombin was labelled using the TMT-126 reagent, and the samples after 1, 2 and 4 hours incubation with ECV by TMT-127. The ratios TMT-127/TMT-126 then reflect changes resulting from proteolysis of intact prothrombin into its activation products. In Figure 1B the labeled lysine residues are ranked in sequence order, both in prothrombin numbering (left), and in chymotrypsin numbering (right). In the bar graph of Figure 1B, a ratio higher

than 1 represents an increase in reactivity, and a ratio below 1 a decrease. The color intensities correspond with the various incubation times with ECV (see legend of Fig. 1). During prothrombin to thrombin conversion, changes in lysine reactivity were limited to a few sites only (Fig. 1B). These changes were time-dependent, and became particularly apparent at the 2 and 4 hour time points, after the lag-phase wherein the inactive derivatives prethrombin-1 and fragment-1 were predominant (Fig. 1A). The most striking differences were detected at the C-terminal side of the protease domain, with a 16-fold increase in reactivity of Lys235+236 and Lys240. Another increase occurred at Lys145, located in the 140-loop. A decrease in reactivity was detected for Lys185+186d, Lys202 and Lys224, which are located in the 180-loop and in the 220-loop or in close proximity thereof. These data show that the conversion of prothrombin into thrombin is accompanied by changes in several of the surface loops of the catalytic domain.

The activation of FX was probed during its activation by the FX-activating protease from Russell's viper venom (RVV-X). SDS/PAGE showed the conversion of FX into FXa over time, as apparent from the loss of the activation peptide after cleavage of the Arg-Ile bond in the heavy chain, leaving Ile16 (in chymotrypsin numbering) representing the new N-terminus of the catalytic domain.[23] After 4 hours an additional band appeared (~47 kDa) as the result of autocatalytic cleavage, without compromising FXa catalytic activity (Fig. 2A).[24] TMT-labelling displayed a slightly different lysine reactivity motif compared to prothrombin activation (Fig. 2B). In contrast to thrombin, the C-terminal end of the FXa protease domain did not show any changes in lysine reactivity, despite the fact that both prothrombin and FX have a conserved lysine residue (Lys236) available as a sensor for changes in this region. Regarding the 220-loop, however, FXa proves to be similar to thrombin in that Lys223+224 display a 4-fold decrease in reactivity (Fig. 2B). Also Lys156 showed a significant, time-dependent decrease. One peptide, containing Lys90+96, suggested increased reactivity, but peptides containing the individual lysine residues did not confirm this change in the 90-loop (Fig. 2B). These data show that FXa is particularly similar to thrombin in that lysine residues in the 220-loop are less reactive after proteolytic activation.

Time-resolved lysine labelling of FIX upon conversion into FIXα and FIXa

Activation of FIX by FXIa follows a sequential mechanism with initial cleavage at Arg145$^{(FIX)}$ followed by a second cleavage at Arg180$^{(FIX)}$, or Arg15 in chymotrypsin numbering.[25] On SDS/PAGE this becomes apparent from the formation of a FIXa heavy chain (HC, 28 kDa) and light chain (LC, 18 kDa) (Fig. 3A). To confirm the generation of the proper N-terminus, the sites of proteolysis were verified by mass spectrometry.

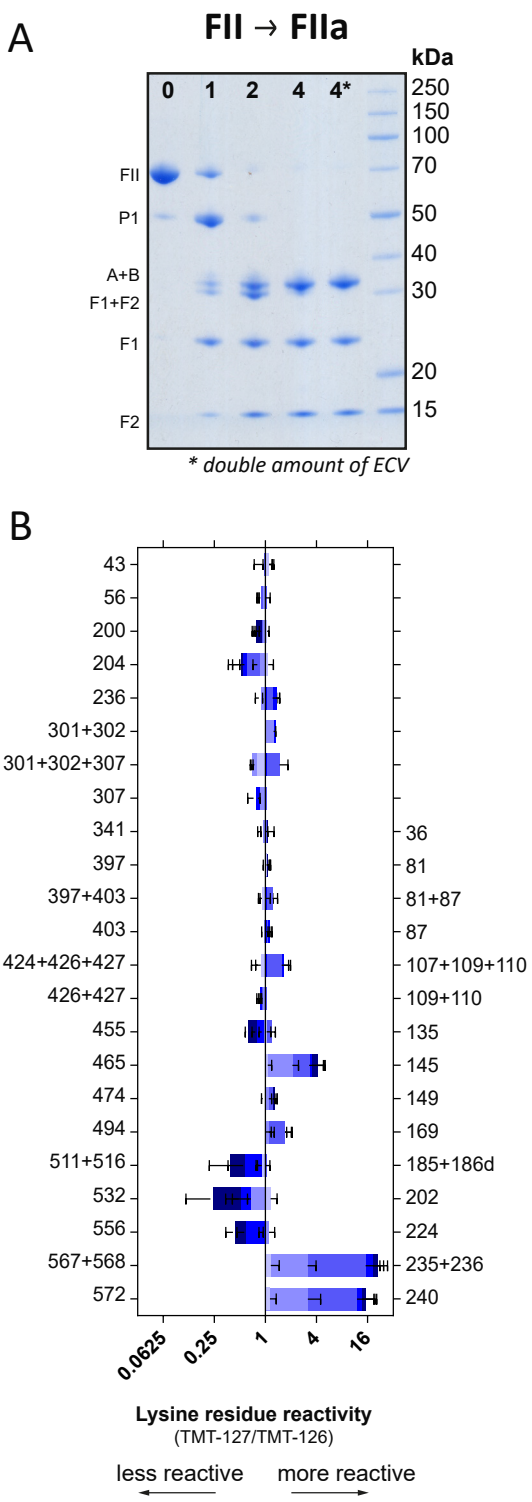


Figure 1: **Prothrombin activation by Echis carinatus snake venom.** About 10 µM of prothrombin was incubated with 0.2 U/mL Echis carinatus venom (ECV) in a buffer containing 20 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$ (pH 7.4). Activation mixtures were incubated with ECV for 1, 2, and 4 hours at 37 °C. An extra sample was incubated for 4 hours with 0.5 U/mL of ECV (indicated with an asterisk). **(A)** Samples of the activation mixture were analysed by 10% SDS PAGE under non-reducing conditions. The various products are indicated at the left, and represent intact prothrombin (FII), prethrombin-1 (P1), thrombin (A+B chain), fragment 1+2 (F1+F2), fragment-1 (F1) and fragment-2 (F2). **(B)** Subsequent to activation with ECV, the same samples were labelled with TMT-127 (time points 1, 2 and 4 hours) or TMT-126 (zero time point) as described in Experimental procedures. TMT-127/TMT-126 ratios are represented in a bar diagram for individual lysine residues after activation for 0 hours (lightest blue), 1 hour (light blue), 2 hours (blue), 4 hours (darker blue) and 4 hours with double amount of ECV (darkest blue). Lysine residues are indicated in prothrombin numbering (left) and chymotrypsin numbering (right).

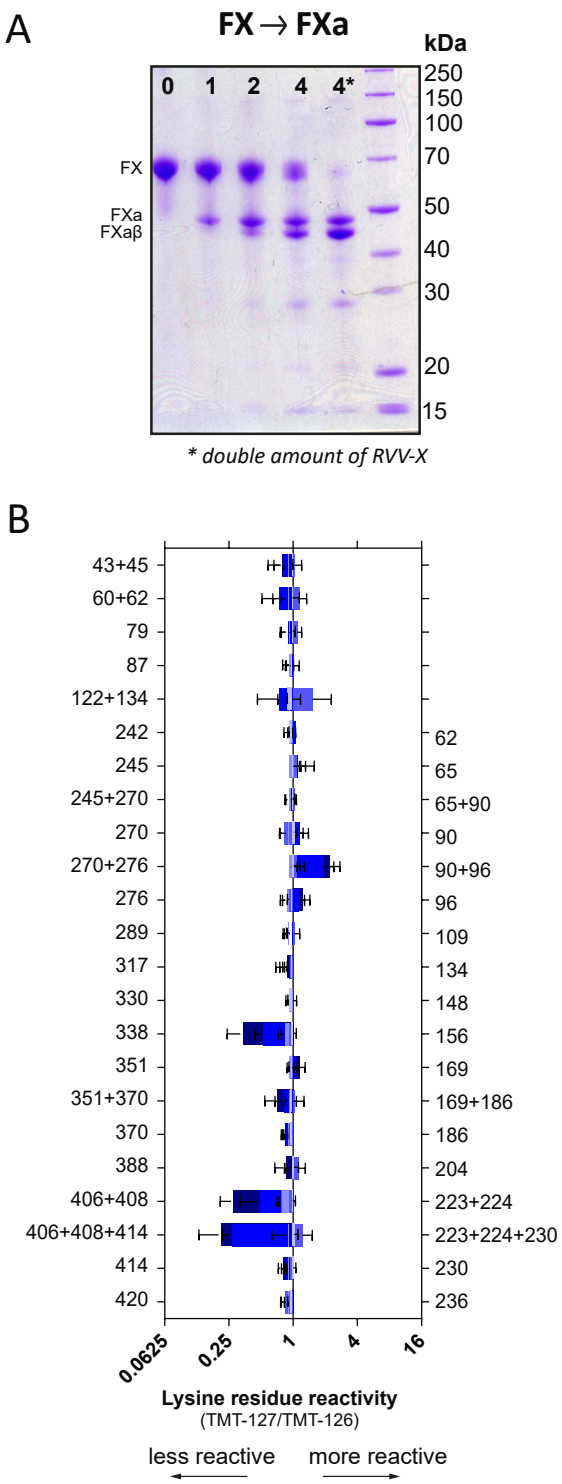


Figure 2: **Activation of FX by Russell's viper venom.** Activation of FX (6 µM) was achieved by incubation with Russell's viper venom protease (RVV-X) (2.26 nM) in buffer containing 20 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$ (pH 7.4). FX was activated for either 1, 2 or 4 hours prior to labelling and MS analysis. An extra sample was incubated for 4 hours with a double amount of RVV-X to ensure full conversion to FXa (indicated with an asterisk). **(A)** Samples of the activation time points were run on a 10% SDS PAGE under non-reducing conditions. The activation products are indicated at the left (intact FX, FXa and FXaβ). **(B)** The same samples were labelled with TMT-127. Extent of lysine residue modification was compared to the labelling of zymogen FX with TMT-126. TMT-127/TMT-126 ratios are given in a bar diagram for the individual lysine residues after activation for 0 hours (lightest blue), 1 hour (light blue), 2 hours (blue), 4 hours (darker blue) and 4 hours with double amount of RVV-X (darkest blue). Lysine numbering is according to FX (left) and chymotrypsin (right).

Figure 3B shows the labelling efficiency of three peptides covering each of the two FIX cleavage sites and, in addition, contain at least one lysine residues. Because of FXIa-mediated proteolysis at the cleavage sites, peptides overlapping the cleavage sites disappear over time in FIXa compared to FIX, reducing the TMT-ratio close to zero (Fig. 3B). This confirms that, in the presence of Ca^{2+}, cleavage occurs at $Arg145^{(FIX)}$ and Arg15, the latter generating Val16 as the novel N-terminus.

Besides FIXa, SDS/PAGE shows formation of an intermediate (FIXα) resulting from a single cleavage at $Arg145^{(FIX)}$ (Fig. 3A). FIXα is represented as a band corresponding to the HC with attached the activation peptide (additional mass of 15 kDa). Previously, it has been demonstrated that this intermediate accumulates upon activation of FIX in the presence of Mn^{2+}.[25,26] Indeed, peptides containing Arg15 and Lys23 did not display a decrease in TMT-ratio (Fig. 3B).

For other lysine-containing peptides, the reactivity towards TMT labels was examined as performed for prothrombin and FX (Fig. 4). Strikingly, conversion of FIX into FIXα was accompanied by negligible changes (Fig. 4A). A minor increase was observed for Lys63, but this proved insignificant due to a high variation in the labelling quantification of this particular peptide. Full activation of FIX to FIXa did display some small effects in lysine reactivity (Fig. 4B). The FIXa 220-loop now showed a 1.2-fold increase in reactivity of Lys222+224, which was absent in FIXα. However, in comparison to prothrombin and FX, FIX proteolysis produced only limited changes. These data indicate that the cleavage at Arg15, resulting in the new N-terminus Val16, is not accompanied by the same changes as those occurring upon activation of FX and prothrombin.

The 220-loop in FIXα and FIXa compared to prothrombin and FX

A closer inspection of the lysine reactivity patterns in Fig. 1, 2 and 4 reveals that the most conspicuous differences occur in the 220-loop. The 220-loops of FIX and FX bear two homologous lysine residues, one of which is also conserved in thrombin (Fig. 5A). Both prothrombin and FX activation initiate a marked, time-dependent decrease in reactivity of the 220-loop lysine residues (Fig. 5B and C). Strikingly, this does not occur upon activation of FIX to FIXα or FIXa (Fig. 5D and E). Although some increase in reactivity was observed in FIXa (Fig. 5E), the extent of these changes was small compared to those seen in thrombin and FXa. These data indicate that the FIX zymogen and FIXa are virtually indistinguishable by our footprinting analysis.

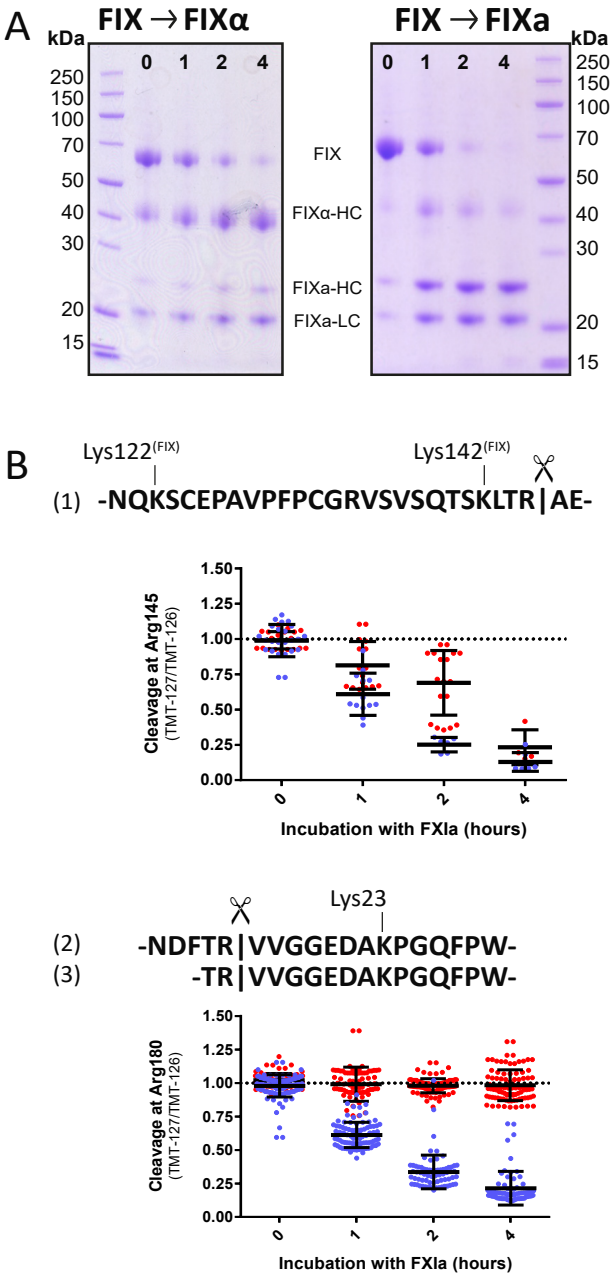


Figure 3: **Activation of FIX by activated factor XI.** FIX (10 µM) was incubated with activated FXI (FXIa, 1.66 nM) in buffer containing 20 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$ (pH 7.4) or, for formation of to FIXα, in the same buffer containing 6.8 mM $MnCl_2$ instead of $CaCl_2$. **(A)** 10% SDS/PAGE analysis under reducing conditions of samples taken after 0, 1, 2 and 4 hours of incubation. The various fragments are FIXα-HC (heavy chain of FIXα), FIXa-HC (heavy chain of FIXa) and FIXa-LC (light chain of FIXa and FIXα). **(B)** Conversion of FIX to FIXa and FIXα was assessed by TMT-labelling. Samples of activated FIX were labelled with TMT-127 and compared with non-activated FIX labelled with TMT-126. Following TMT-labelling, proteolytic digestion and mass spectrometry, peptides that cover the two cleavage sites (Arg145$^{(FIX)}$-Ala146$^{(FIX)}$ and Arg15-Val16) were identified. Due to cleavage in these sites, these peptides gradually disappeared. This drove the TMT-127/TMT-126 ratio of the lysine residues towards zero. Activation in presence of Ca^{2+} (blue dots), shows a reduction of TMT-ratios of peptides that cover both cleavage sites. Activation in presence of Mn^{2+} (red dots) shows a reduction of peptides that cover cleavage site Arg145$^{(FIX)}$-Ala146$^{(FIX)}$, while the ratio of the peptide covering Arg15-Val16 remains ~1.

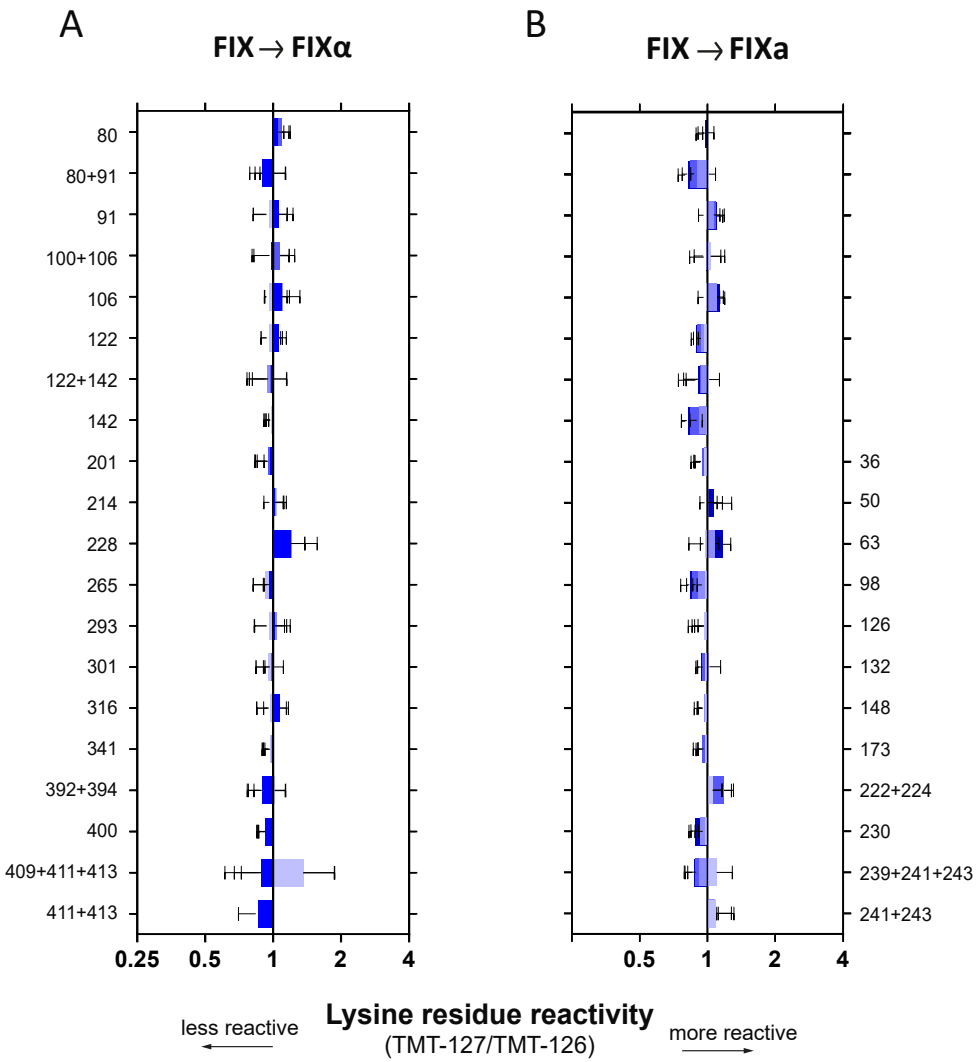


Figure 4: **Activation of FIX by activated factor XI.** FIX was activated by FXIa under the conditions described in Fig. 3, in the presence of **(A)** Ca^{2+} or **(B)** Mn^{2+}. Activated FIX was labelled with TMT-127 and compared to zymogen FIX, labelled with TMT-126. TMT-127/TMT-126 ratios are given in a bar diagram for the individual residues after activation for 0 hours (lightest blue), 1 hour (light blue), 2 hours (blue), 4 hours (darkest blue). TMT-ratios of lysine residues shown were based on peptides that do not cover one of the two cleavage sites (see Fig. 3). Lysine residues are given in FIX (left) and chymotrypsin (right) numbering.

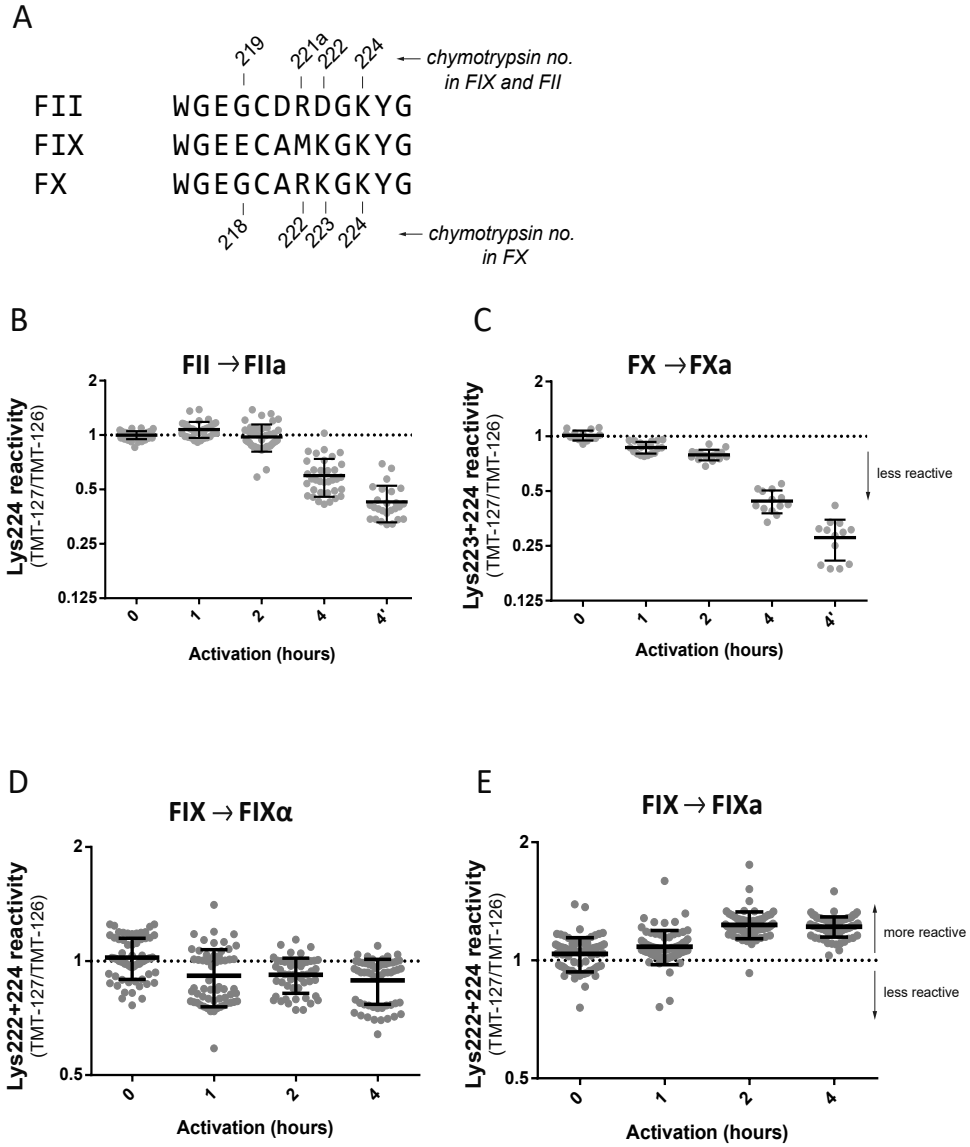


Figure 5: **Lysine reactivity of the 220-loop upon activation of prothrombin, FX and FIX. (A)** Alignment of 220-loops of prothrombin (FII), FX and FIX. The chymotrypsin numbering in this loop is not consistent for these proteases. Therefore, the numbering in FIX and prothrombin is given above the alignment and the numbering of FX at the bottom. The lysine reactivity of the 220-loop was assessed by TMT-labelling following activation of **(B)** prothrombin, **(C)** FX and **(D)** FIX in presence of Ca^{2+} and **(E)** FIX in presence of Mn^{2+}. Samples from the incubation mixtures were labelled by TMT-127 and compared to the zero time point representing the corresponding non-activated zymogen labelled by TMT-126. Each dot represents the ratio between labelling intensities (TMT-127/TMT-126) of peptides that contain either Lys224 or Lys224 and Lys222.

Exposure of protease domain N-terminus and 220-loop

So far, we assessed the relation between lysine reactivity and limited proteolysis at the previously established sites. According to the common paradigm, the new N-terminus of the catalytic domain should insert into the active site. However, the notion that in FVIIa this only occurs after interaction with tissue factor or after binding of an irreversible inhibitor into the active site,[16,17,27] prompted us to explore this in FIXa, FXa and thrombin. We further focused on the 220-loop because the lysines therein proved different between FIXa on the one hand, and FXa and thrombin on the other hand (Fig. 5). For this purpose, N-terminus labelling of FIXa, FXa and thrombin was assessed prior and after reaction with an irreversible inhibitor, or pseudo-substrate. The inhibited derivatives were indicated as FIXa-EGR, FXa-FPR and thrombin-FPR,

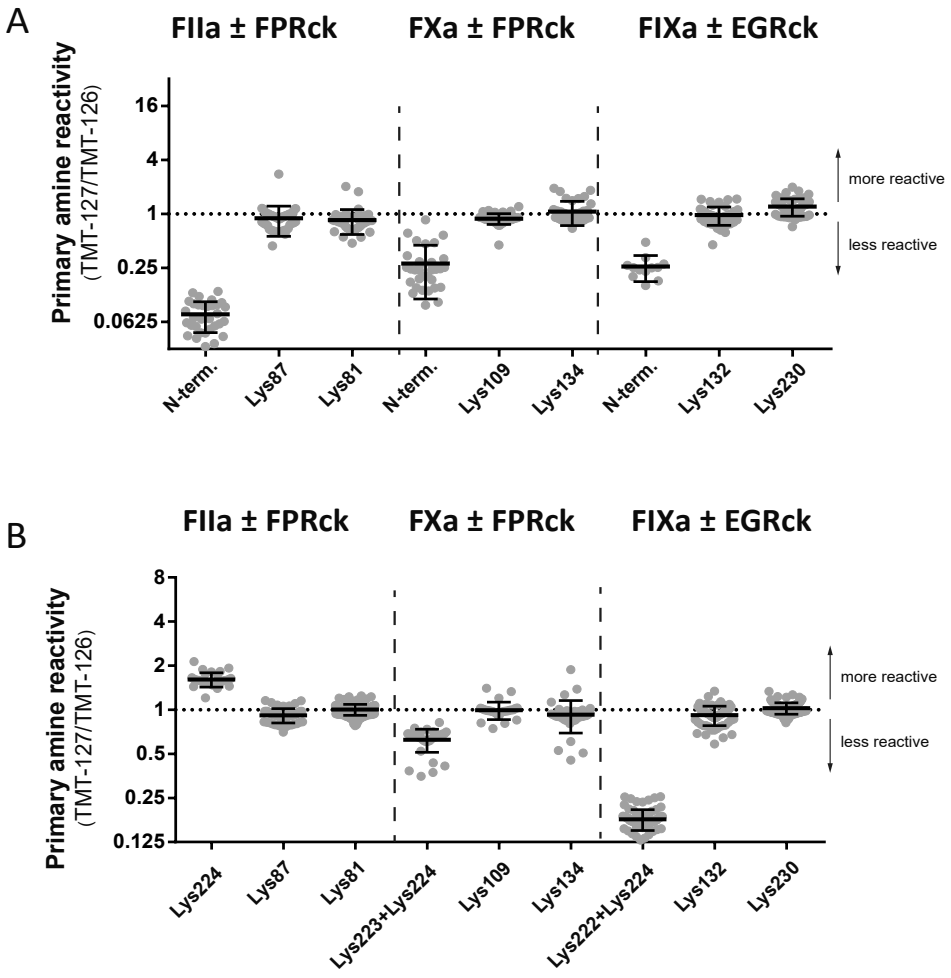


Figure 6 *(figure legend on next page)*

Figure 6: T**MT-labelling of the N-terminus and 220-loop lysine residues following binding of a pseudo-substrate.** FXa (4 µM) and thrombin (FII) (2 µM) were incubated with an excess (10 mM) of H-D-Phe-Pro-Arg-chloromethylketone (FPRck) in a buffer containing 20 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$ (pH 7.4) for 45 min at 37 °C. Similarly, FIXa (4 µM) was inactivated with H-Glu-Gly-Arg-chloromethylketone (EGRck). The reaction was stopped by removal of ERGck and FPRck using Zeba Spin Columns according to the manufacturer's instructions. **(A)** After TMT-labelling and proteolytic digestion, peptides of the N-terminus were identified from the MS^2 spectra employing a database search as described in Experimental procedures. TMT label intensities were inferred from MS^3 fragmentation (higher energy collision dissociation) on peptide fragments obtained from the MS^2 fragmentation (collision induced dissociation). This served to prevent interference of TMT-labelled lysine residue on position 23 in FIXa and FXa of peptides covering the N-terminus, V(+TMT)VGGEDAK(+TMT)PGQFPW and I(+TMT) VGGQECK(+TMT)DGECPW, respectively. Only the fragments (b^+-ions) that correspond to peptide part V(+TMT)VGGED of FIXa and peptide part I(+TMT)VGGQE of FXa were selected for further quantification of TMT-ratios in the MS^3 spectra. TMT-127/TMT-126 ratios were inferred from the MS^3 spectra according to peak intensities of masses 126.127 and 127.131 (Xcalibur 2.0). The same method of quantification was employed on control residues Lys132 and Lys230 for FIXa, Lys109 and Lys134 for FXa and Lys87 and Lys81 for thrombin using y- and b-ions containing these lysine residues. Each dot represents a measurement of a MS^3 fragment with corresponding TMT-ratio. **(B)** Quantification of the complete lysine residue set, including Lys222 and Lys224, was performed following MS^2 mass spectrometry using a Thermo Orbitrap XL mass spectrometer and Proteome Discoverer software as described previously.[21]

respectively. Because the N-terminal peptides in FIXa and FXa, in addition to the amine of the N-terminus, also contain a lysine residue, additional fragmentation was required to ensure that only the labeled N-terminus was quantified by mass spectrometry (Fig. 6). When comparing thrombin to thrombin-FPR, labelling of the N-terminus was about 10-fold lower in thrombin-FPR than in thrombin (Fig. 6A). Analysis of arbitrary chosen lysine residues, taken as internal controls, did not indicate any significant change. This indicates that the observed decrease is specific for the N-terminus of thrombin. Remarkably, the reduction in N-terminus labelling was accompanied by a small increase in reactivity of Lys224 in the 220-loop (Fig. 6B). The N-terminus of the FXa protease domain showed a smaller decrease in reactivity in comparison with thrombin, as a 4-fold decrease was observed upon reaction with FPRck (Fig. 6A). This was accompanied by a slight decrease in reactivity for the lysine residues of the 220-loop (Fig. 6B). The difference between FIXa and FIXa-EGR, however, proved more prominent. While the N-terminus displayed an equal extent of protection as FXa, the 220-loop lysine residues showed a 6-fold decrease in reactivity (Figs. 6A and B). In contrast to FIXa-EGR, the incorporation of a pseudo-substrate into FXa and thrombin did not affect the lysine residues in the 220-loop. These data show that FIXa differs from FXa and thrombin in that its changes in the 220-loop occur upon interaction with an irreversible inhibitor only, while in FXa and thrombin proteolytic cleavage alone seems sufficient to induce these changes.

DISCUSSION

In this study we set out to probe structural changes that occur upon activation of coagulation factors prothrombin, FX and FIX. Therefore, we investigated the reactivity of primary amines, in particular lysine residues, towards small molecular probes. The use of isobaric TMTs together with tandem mass spectrometry enabled us to quantify labelling of lysine residues upon zymogen to protease conversion (Fig. 1, 2 and 4). Previously, we have used TMTs to probe strong interactions, for instance to identify binding interfaces between LDL receptor-related protein-1 (LRP) and RAP or to assess protein integrity within FVIII variants.[18,19,21] The TMT approach compares the labelling efficiency at particular amine groups under two distinct conditions. In the current study we employed this approach to include samples from coagulation proenzymes with different activation times, and enzymes with *versus* without inhibitor to fill the S1-pocket. It should be noted that TMT footprinting data are expressed in ratios and thus provide no absolute quantification. In our previous studies, TMT ratios between 1.5 and 2 were observed for strong interactions.[18,19,21] In the present study we observed much higher ratios. This is mainly due to a difference in labelling conditions. In preliminary experiments (data not shown), we observed that ratios increased upon shortening labelling time or lowering temperature. Thus, labelling kinetics strongly contribute to the dynamic range of TMT footprinting as a tool for probing intramolecular changes. Lysine labelling proved particularly useful in that it could be performed in solution, while activation was ongoing. A limitation, however, is that we were unable to analyze the entire composition of the coagulation factors by a single protocol. Due to their abundance in negatively charged residues, the γ-carboxy-glutamic acid containing domains were not covered. In addition, glycan-rich areas such as the FIX activation peptide could not be analysed. Studying these regions requires specific mass spectrometry approaches.[28,29] Nevertheless, we achieved sequence coverage of the major part of the coagulation factors in our study and in particular of the catalytic domains.

Prominent differences in lysine reactivity were observed between the prothrombin, FX, FIX and the activation products thereof (Fig. 1, 2 and 4). Lysine residues that displayed a remarkably large increase in labelling were found at the C-terminal end of prothrombin (Fig. 1). As part of the exosite II, the C-terminal lysine residues have been described to bind the kringle-2 domain in fragment-2 with high affinity in order to prevent premature interaction with ligands such as glycoprotein Ibα and heparin.[30,31] Uncovering of exosite II, with concomitant increase in lysine residue exposure, would explain the observed increase in labelling.

Besides exosite II, conversion of prothrombin to thrombin primarily displays changes in the protease domain surface loops. The 140-, 180- and 220-loops all

indicate a change upon activation (Fig. 1). These loops are adjacent to the N-terminus insertion site and have previously been described as 'activation loops' showing intense reordering upon activation and ligand binding by X-ray crystallography, NMR and HDX techniques.[32-37] The notion that prothrombin surface loops are highly sensitive to activation is supported by our time-resolved lysine labelling. In comparison to prothrombin, FX activation displayed fewer changes. Obviously, the kringle-2 interaction and the particularly large 180-loop of prothrombin are not conserved in FX. Moreover, the 140-loop did not display any change upon activation (Fig. 2). As for the 220-loop, activation of both FX and prothrombin resulted in a decrease in lysine reactivity. This suggests that these zymogens share, at least partially, the same 'activation loop' mechanism (Fig. 1, 2 and 5). In this respect, FIX is different, as its catalytic domain seems indistinguishable before and after limited proteolysis at the activation sites (Fig. 4). Assuming that zymogen-to-protease conversion can be expressed in the reorientation and/or stabilization of surface loops of the protease domain, the absence of substantial changes in FIX indicates that FIXa is more zymogen-like than FXa and thrombin.

Surprisingly, binding of a substrate analogue to the active site of FIXa also affected reactivity of the lysine residues in the 220-loop (Fig. 6B). Thereby, the lysine reactivity of the FIX 220-loop appears to mirror the reactivity in FX and prothrombin. In fact, while activation of FX and prothrombin showed a strong response in the 220-loop reactivity, substrate binding only had limited effect (Fig. 5 and 6B). FIX displayed the opposite, activation left the 220-loop almost unaffected while substrate binding yielded a major change (Fig. 5B and 6B). Although the 220-loops are highly homologous between the studied coagulation factors some key differences are present (Fig. 5A). When surveying the crystal structures of prothrombin, a flexible and open orientation of Lys224 and Glu217 emerges compared to a stable, close interaction in thrombin (Fig. 7A and B).[38-41] In line with the structural information, a decrease in reactivity of Lys224 suggests that the Lys224-Glu217 interaction is formed (Fig. 1B). In FIX, the 220-loop contains an extra glutamic acid at position 219 which, in conjunction with Glu217 could interact with Lys224.[42] Furthermore, FIXa crystal structures display an interaction between Lys222 with Glu186, not present in FXa or thrombin. Lysine labelling indicates that at least one of these interactions is diminished in FIXa-EGR (Fig. 6).

In the allostery model for trypsin-like proteases proposed by Di Cera and co-workers, the 215-217 segment (opposite to the 220-loop) plays a pivotal role in blockage or opening of the active site.[43] Our labelling studies together with the available structural information might indicate that prothrombin activation can directly affect the 215-217 segment by interaction between Lys224 and Glu217 (Fig. 7). Thereby, the decrease in Lys224 reactivity reflects a transformation of prothrombin with a

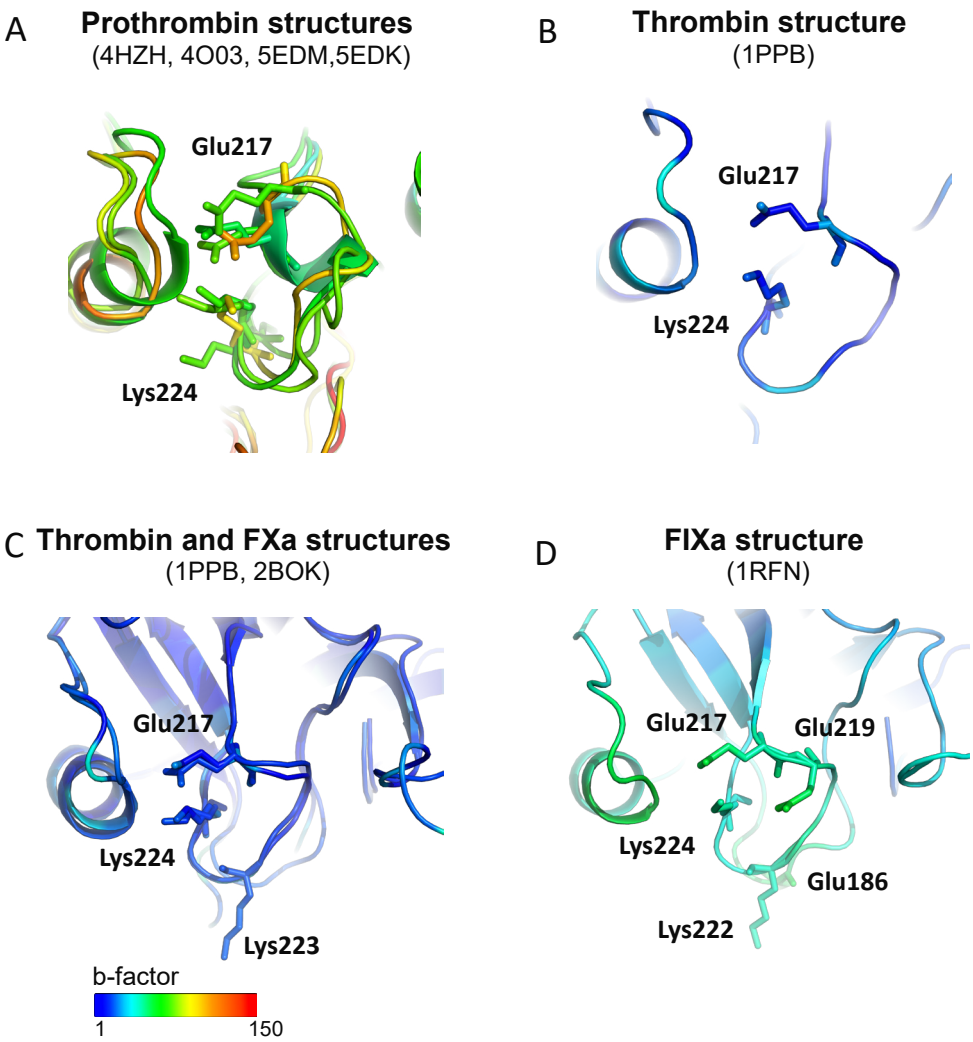


Figure 7: **Structural information on the 220-loop of prothrombin, thrombin, FXa and FIXa.** Structures represent **(A)** prothrombin, **(B)** thrombin, **(C)** FXa and **(D)** FIXa.[38-42] Lysine residues and glutamic acids that might interact are indicated by stick representations. Images were created using crystallography data deposited in the Protein Data Bank, PDB codes are given per image. Colour of the backbone structure represents the B-factor. Images were made in PyMol supplied by Schrödinger (Mannheim, Germany).

collapsed active site (Z*) to thrombin with an open active site (E).[43] FIXa appears to be different in this respect. By lack of interaction between Lys224 and Glu217/Glu219, the 215-217 segment might remain unaffected and could still obstruct the active site entry (going from Z* to E*). Unfortunately, we are unable determine the separate contributions of Lys224 and Lys222 to accurately specify the interaction of Lys224. Alternative to the trypsin-like allostery, our data could be viewed in an ensemble-defined allostery for coagulation proteases.[16] In this hypothesis a sustained flexibility of the protease domain surface loops could form the basis of the poor activity in FIXa. By an overall lack of interactions within the 220-loop (either Lys222 or Lys224), the FIXa protease domain may not be driven towards an active conformation. Instead, the protease domain may occupy transition states that are predominantly zymogen-like (Fig. 4).

Another finding in our study relates to the decrease in N-terminus labelling that occurred upon filling the S1-pocket with a substrate analogue (Fig. 6A). In activated coagulation factor VII (FVIIa), impaired N-terminal insertion has been described as a reason for its poor activity.[16,44] In FVIIa, insertion of its N-terminus is greatly facilitated by presence of an inhibitor in the active site or by interaction with its natural cofactor tissue factor.[16,17,27,44] In this regard it seems surprising that the most prominent effect of filling the substrate binding site was observed for thrombin. Earlier studies have observed dynamic changes in the thrombin N-terminus as detected by various NMR techniques and HDX.[35-37] This suggests that the N-terminus requires some freedom in switching between inserted and free. The presence of catalysis-enhancing ligands such as hirugen has been reported to rigidify the N-terminus,[35] perhaps by driving the equilibrium toward the inserted form. In our study, the change in reactivity of the N-terminus might (Fig. 6A) reflect that proteolytic activation alone is not sufficient to drive thrombin into its optimal protease-like conformation.

It seems striking that in the current comparison between FIXa, FXa and thrombin, FIXa proves to be the most zymogen-like. As such, it more resemble FVIIa. From a regulatory perspective, it might be beneficial that enzymes upstream in the coagulation cascade have the lowest intrinsic catalytic activity. Perhaps, nature developed a more specialized control system (i.e. cofactors) for these factors that overtakes the (Na^+-dependent) activation mechanism found in thrombin.[45] Limiting the risk of premature activation, followed by amplification thereof downstream in the cascade, may be more important for FVIIa and FIXa than for thrombin.

EXPERIMENTAL PROCEDURES

Nomenclature

To facilitate comparison between protease domains, chymotrypsin numbering is used throughout this paper. In case of residues that are outside of the protease domain, protein numbering is used with the abbreviated coagulation protease given in brackets in superscript, as in Arg145$^{(FIX)}$.

Materials

Precast SDS/PAGE gels were obtained from Invitrogen (Breda, the Netherlands). The TMTduplex Isobaric Tagging kit, Zeba Spin columns (7K, 0.5 mL) and endoproteases chymotrypsin and Glu-C were from Thermo Scientific (Breda, the Netherlands). The endoprotease Arg-C was supplied by Promega (Leiden, the Netherlands). Pseudo-substrates H-D-Phe-Pro-Arg-chloromethylketone (FPRck) and H-Glu-Gly-Arg-chloromethylketone (EGRck) were from Bachem (Bubendorf, Switzerland). Software programs Xcalibur 2.0 and Proteome Discoverer were supplied by Thermo Scientific. Image of crystal structures were processed in PyMol supplied by Schrödinger (Mannheim, Germany).

Proteins

Human FIX, FX and prothrombin were isolated from plasma as described previously.[26,46,47] Conversion of prothrombin, FX and FIX to their activated forms and subsequent purification has been described previously.[26,47,48] Proteins were quantified by the Bradford method using human serum albumin as standard.[49] Snake venom activators from *Echis carinatus* (ECV) and Russell's viper (RVV-X) were supplied by Pentapharm (Aesch, Switzerland). FXIa was purchased from Stago (Leiden, the Netherlands).

TMT-labelling

To monitor activation of FIX, FX and prothrombin, TMT-labelling was performed directly after activation in buffer containing 20 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$ (pH 7.4) and 2.5 mM of TMT-labels (dissolved in DMSO). When FIX was activated in the presence of Mn^{2+}, $CaCl_2$ was replaced by 6.8 mM of $MnCl_2$. Labelling of the activated (TMT-127) and the non-activated coagulation factors (TMT-126) was initiated by transfer of FIX(a), FX(a) or (pro)thrombin (0.45, 0.39 or 0.36 μM, respectively) to the reaction buffer containing TMTs. For comparison of the pseudo-substrate inhibited forms of FIXa, FXa and thrombin versus the free enzymes, TMT-labelling was performed by incubation of 0.5 μM enzyme (inhibited or free) with 2.5 mM TMT-labels in 20 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$ (pH 7.4). Incubation with

TMT labels proceeded for 7.5 min at 25 °C under stirring at 650 rpm. The reactions were quenched by addition of 0.75% hydroxylamine for 15 min at room temperature. Subsequently, equimolar amounts of protein were combined from the TMT-126 and TMT-127 containing reaction mixtures. TMT-126 was used for zero time points and for the non-inhibited proteases, whereas TMT-127 was used for the 1, 2 and 4 hour time points and for the inhibited form of the proteases. The combined mixtures were then processed for mass spectrometry analysis as described.[21] In short, proteins were reduced by incubation at 56 °C for 45 minutes with 7.7 mM dithiothreitol in 100 mM ammonium bicarbonate. Subsequently, free cysteines were acetylated by addition of 2-iodoacetamide (10.5 mM final concentration) in 100 mM ammonium bicarbonate at room temperature for 30 minutes in the dark. Peptides were obtained using chymotrypsin and Glu-C (V8-protease) endoproteases for FIX, FX and prothrombin. For FX, additional peptides were generated using Arg-C. Peptides were desalted using a C18 Ziptip (Merck Millipore Corporation, Amsterdam, the Netherlands). Digestion and desalting was performed according to the manufacturer's instructions.

Mass Spectrometry

Identification and quantification of TMT-labels from peptides of the zymogen activation studies were performed using a Thermo Scientific LTQ OrbitrapXL mass spectrometer as described elsewhere.[21] A Thermo Scientific Orbitrap Fusion Tribrid mass spectrometer was used for the assessment of TMT-labeled N-termini. Briefly, peptides were separated by reverse-phase chromatography on an in-house packed C18 column. Samples were loaded for 17 minutes at 250 nL/min in 95% solution A (0.5% acetic acid) and 5% solution B (0.5% acetic acid and 80% acetonitrile). After equilibration for 5 min in the same mixture (17-22 minutes), peptides were eluted by increasing solution B from 5% to 30% (22-132 min) and 30% to 60% (132-147 minutes), followed by a 10-minute wash to 90% and 5-min regeneration to 5%. Using a nano-electrospray ion source set at 2.1 kV, peptides were sprayed into the Orbitrap Fusion mass spectrometer. The five most intense precursor ions in the full scan (400-1500 m/z, 120.000 resolving power) were fragmented using Collision Induced Dissociation (CID) in the ion trap with a maximum injection time of 60 ms and 35% collision energy. The MS^2 fragments were analysed in the ion trap (300-2000 m/z). CID-derived spectra were used for peptide identification with PEAKS studio 7.0 (Bioinformatics Solutions Inc, Waterloo, Canada), employing a database search against the human proteome (uniprot: organism 9606 and keyword kw 0181) with optional TMT modification of lysine residues and the N-terminus. The five most intense MS^2 ions were subjected to MS^3 fragmentation using Higher-energy Collisional Dissociation (HCD) in the ion trap with a maximum injection time of 50 ms and 65% collision energy. Resulting fragments were analysed in the Orbitrap (100-500 m/z, 60000 resolution power).

REFERENCES

1. Davidson CJ, Hirt RP, Lal K, Snell P, Elgar G, Tuddenham EG, McVey JH. Molecular evolution of the vertebrate blood coagulation network. *Thromb Haemost.* 2003 **89** 420-428.
2. Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry.* 1991 **30** 10363-10370.
3. Page MJ, Di Cera E. Serine peptidases: classification, structure and function. *Cell Mol Life Sci.* 2008 **65** 1220-1236.
4. Hedstrom L. Serine protease mechanism and specificity. *Chem Rev.* 2002 **102** 4501-4524.
5. Rawlings ND, Barrett AJ. Families of serine peptidases. *Methods Enzymol.* 1994 **244** 19-61.
6. Huntington JA. Thrombin plasticity. *Biochim Biophys Acta.* 2012 **1824** 246-52.
7. Bjelke JR, Olsen OH, Fodje M, Svensson LA, Bang S, Bolt G, Kragelund BB, Persson E. Mechanism of the Ca^{2+}-induced enhancement of the intrinsic factor VIIa activity. *J Biol Chem.* 2008 **283** 25863-25870.
8. Mathur A, Zhong D, Sabharwal AK, Smith KJ, Bajaj SP. Interaction of factor IXa with factor VIIIa. Effects of protease domain Ca^{2+} binding site, proteolysis in the autolysis loop, phospholipid, and factor X. *J Biol Chem.* 1997 **272** 23418-23426.
9. Rezaie AR, Esmon CT. Asp-70-->Lys mutant of factor X lacks high affinity Ca^{2+} binding site yet retains function. *J Biol Chem.* 1994 **269** 21495-21499.
10. Pineda AO, Carrell CJ, Bush LA, Prasad S, Caccia S, Chen ZW, Mathews FS, Di Cera E. Molecular dissection of Na^{+} binding to thrombin. *J Biol Chem.* 2004 **279** 31842-31853.
11. Schmidt AE, Stewart JE, Mathur A, Krishnaswamy S, Bajaj SP. Na^{+} site in blood coagulation factor IXa: effect on catalysis and factor VIIIa binding. *J Mol Biol.* 2005 **350** 78-91.
12. Underwood MC, Zhong D, Mathur A, Heyduk T, Bajaj SP. Thermodynamic linkage between the S1 site, the Na^{+} site, and the Ca^{2+} site in the protease domain of human coagulation factor Xa. Studies on catalytic efficiency and inhibitor binding. *J Biol Chem*. 2000 **275** 36876-36884.
13. Lechtenberg BC, Freund SM, Huntington JA. An ensemble view of thrombin allostery. *Biol Chem.* 2012 **393** 889-898.
14. Krishnaswamy S. The transition of prothrombin to thrombin. *J Thromb Haemost.* 2013 **11** 265-276.
15. Vogt AD, Chakraborty P, Di Cera E. Kinetic dissection of the pre-existing conformational equilibrium in the trypsin fold. *J Biol Chem.* 2015 **290** 22435-22445.
16. Higashi S, Matsumoto N, Iwanaga S. Molecular mechanism of tissue factor-

mediated acceleration of factor VIIa activity. *J Biol Chem.* 1996 **271** 26569-26574.
17. Rand KD, Jørgensen TJ, Olsen OH, Persson E, Jensen ON, Stennicke HR, Andersen MD. Allosteric activation of coagulation factor VIIa visualized by hydrogen exchange. *J Biol Chem.* 2006 **281** 23018-23024.
18. Ebberink EH, Bouwens EA, Bloem E, Boon-Spijker M, van den Biggelaar M, Voorberg J, Meijer AB, Mertens K. Factor VIII/V C-domain swaps reveal discrete C-domain roles in factor VIII function and intracellular trafficking. *Haematologica.* 2017 **102** 686-694.
19. Bloem E, Meems H, van den Biggelaar M, van der Zwaan C, Mertens K, Meijer AB. Mass spectrometry-assisted study reveals that lysine residues 1967 and 1968 have opposite contribution to stability of activated factor VIII. *J Biol Chem.* 2012 **287** 5775-5783.
20. Castro-Núñez L, Bloem E, Boon-Spijker MG, van der Zwaan C, van den Biggelaar M, Mertens K, Meijer AB. Distinct roles of Ser-764 and Lys-773 at the N terminus of von Willebrand factor in complex assembly with coagulation factor VIII. *J Biol Chem.* 2013 **288** 393-400.
21. Bloem E, Ebberink EH, van den Biggelaar M, van der Zwaan C, Mertens K, Meijer AB. A novel chemical footprinting approach identifies critical lysine residues involved in the binding of receptor-associated protein to cluster II of LDL receptor-related protein. *Biochem J.* 2015 **468** 65-72.
22. Franza BR Jr, Aronson DL, Finlayson JS. Activation of human prothrombin by a procoagulant fraction from the venom of Echis carinatus. Identification of a high molecular weight intermediate with thrombin activity. *J Biol Chem.* 1975 **250** 7057-7068.
23. Di Scipio RG, Hermodson MA, Davie EW. Activation of human factor X (Stuart factor) by a protease from Russell's viper venom. *Biochemistry*. 1977 **16** 5253-5260.
24. Jesty J, Spencer AK, Nakashima Y, Nemerson Y, Konigsberg W. The activation of coagulation factor X. Identity of cleavage sites in the alternative activation pathways and characterization of the COOH-terminal peptide. *J Biol Chem.* 1975 **250** 4497-4504.
25. Di Scipio RG, Kurachi K, Davie EW. Activation of human factor IX (Christmas factor). *J Clin Invest.* 1978 **61** 1528-1538.
26. Lenting PJ, ter Maat H, Clijsters PP, Donath MJ, van Mourik JA, Mertens K. Cleavage at arginine 145 in human blood coagulation factor IX converts the zymogen into a factor VIII binding enzyme. *J Biol Chem.* 1995 **270** 14884-14890.
27. Madsen JJ, Persson E, Olsen OH. Tissue factor activates allosteric networks in factor VIIa through structural and dynamic changes. *J Thromb Haemost.* 2015 **13** 262-267.

28. Ramström M, Sandberg H. Characterization of γ-carboxylated tryptic peptides by collision-induced dissociation and electron transfer dissociation mass spectrometry. *Eur J Mass Spectrom.* 2011 **17** 497-506.
29. Banazadeh A, Veillon L, Wooding KM, Zabet-Moghaddam M, Mechref Y. Recent advances in mass spectrometric analysis of glycoproteins. *Electrophoresis.* 2017 **38** 162-189.
30. Lechtenberg BC, Freund SM, Huntington JA. GpIbα interacts exclusively with exosite II of thrombin. *J Mol Biol.* 2014 **426** 881-893.
31. Carter WJ, Cama E, Huntington JA. Crystal structure of thrombin bound to heparin. *J Biol Chem.* 2005 **280** 2745-2749.
32. Huntington JA. Slow thrombin is zymogen-like. *J Thromb Haemost.* 2009 **7** 159-164.
33. Malovichko MV, Sabo TM, Maurer MC. Ligand binding to anion-binding exosites regulates conformational properties of thrombin. *J Biol Chem.* 2013 **288** 8667-8678.
34. Croy CH, Koeppe JR, Bergqvist S, Komives EA. Allosteric changes in solvent accessibility observed in thrombin upon active site occupation. *Biochemistry.* 2004 **43** 5246-5255
35. Lechtenberg BC, Johnson DJ, Freund SM, Huntington JA. NMR resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc Natl Acad Sci U S A.* 2010 **107** 14087-14092.
36. Handley LD, Treuheit NA, Venkatesh VJ, Komives EA. Thrombomodulin Binding Selects the Catalytically Active Form of Thrombin. *Biochemistry.* 2015 **54** 6650-6658.
37. Handley LD, Fuglestad B, Stearns K, Tonelli M, Fenwick RB, Markwick PR, Komives EA. NMR reveals a dynamic allosteric pathway in thrombin. *Sci Rep.* 2017 **7** 39575.
38. Pozzi N, Chen Z, Gohara DW, Niu W, Heyduk T, Di Cera E. Crystal structure of prothrombin reveals conformational flexibility and mechanism of activation. *J Biol Chem.* 2013 **288** 22734-22744.
39. Pozzi N, Chen Z, Pelc LA, Shropshire DB, Di Cera E. The linker connecting the two kringles plays a key role in prothrombin activation. *Proc Natl Acad Sci U S A.* 2014 **111** 7630-7635.
40. Pozzi N, Chen Z, Di Cera E. How the Linker Connecting the Two Kringles Influences Activation and Conformational Plasticity of Prothrombin. *J Biol Chem.* 2016 **291** 6071-6082.
41. Bode W, Mayr I, Baumann U, Huber R, Stone SR, Hofsteenge J. The refined 1.9 Å crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* 1989 **8** 3467-3475.

42. Hopfner KP, Lang A, Karcher A, Sichler K, Kopetzki E, Brandstetter H, Huber R, Bode W, Engh RA. Coagulation factor IXa: the relaxed conformation of Tyr99 blocks substrate binding. *Structure.* 1999 **7** 989-996.
43. Gohara DW, Di Cera E. Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends Biotechnol.* 2011 **29** 577-585.
44. Higashi S, Nishimura H, Aita K, Iwanaga S. Identification of regions of bovine factor VII essential for binding to tissue factor. *J Biol Chem.* 1994 **269** 18891-18898
45. Page MJ, Di Cera E. Is Na^+ a coagulation factor? *Thromb Haemost.* 2006 **95** 920-921.
46. Hendrix H, Lindhout T, Mertens K, Engels W, Hemker HC. Activation of human prothrombin by stoichiometric levels of staphylocoagulase. *J Biol Chem.* 1983 **258** 3637-3644.
47. Mertens K, van Wijngaarden A, Bertina RM. The role of factor VIII in the activation of human blood coagulation factor X by activated factor IX. *Thromb Haemost.* 1985 **54** 654-660.
48. Mertens K, Bertina RM. Activation of human coagulation factor VIII by activated factor X, the common product of the intrinsic and the extrinsic pathway of blood coagulation. *Thromb Haemost.* 1982 **47** 96-100.
49. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976 **72** 248-254

6

The 220-loop in coagulation factor IX revisited: role of basic and acidic residues in catalytic activity

Eduard H.T.M. Ebberink[1], Nadia Freato[1], Floris P.J. van Alphen[1], Mariëtte G. Boon-Spijker[1], Maartje van den Biggelaar[1,2], Alexander B. Meijer[1,2] and Koen Mertens[1,2]

Manuscript in preparation

[1]Department of Plasma Proteins, Sanquin Research, 1066 CX Amsterdam, The Netherlands
[2]Pharmaceutical Sciences, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands

ABSTRACT

Like in other serine proteases, proteolytic activation of coagulation factor IX (FIX) involves the generation of a novel N-terminus and intramolecular rearrangements in its catalytic active site. However, structural events that drive rate enhancement of activated FIX (FIXa) have so far remained poorly understood. We previously showed that in the homologous coagulation enzymes factor Xa and thrombin, surface loops in the protease domain, in particular the 220-loop, change upon the conversion of zymogen into the active enzyme. We now investigated the involvement of charged residues in the 220-loop of FIXa, by analyzing the substitution variants E217A, E219A, K222A and K224A. All mutants displayed an impaired activity towards a tripeptide substrate and the natural substrate factor X. The structural implications of these substitutions were assessed using chemical footprinting and mass spectrometry techniques, with particular reference to changes in the 220-loop and exposure of the N-terminus. Primary amine labelling revealed that the insertion of the N-terminus was reduced for substitution variants at 217, 222 and 224, while the opposite was found for the variant E219A. Furthermore, footprinting suggests an intricate relation between N-terminus insertion and the solvent exposure of 220-loop lysines. Conformational changes in the 220-loop were confirmed by hydrogen-deuterium exchange, and were most prominent after inhibition of FIXa by irreversible binding of a pseudo-substrate. We propose that substrate-driven rigidification of the 220-loop, in combination with a more efficient N-terminus insertion provides an important allosteric mechanism in FIX.

INTRODUCTION

Blood coagulation involves a cascade comprising various thrombin-like serine proteases and their respective cofactors.[1,2] In the middle phase of the cascade, coagulation factor IX (FIX) is an essential component in the activation of factor X (FX), as part of a positive feedback loop in the coagulation cascade.[1,2] The importance of FIX is reflected by the bleeding disorder caused by its deficiency or dysfunction, known as *Haemophilia B*.[3]

Serine proteases of the coagulation cascade like thrombin and activated IX (FIXa) share the catalytic domain structure and activation mechanism with other members of the chymotrypsin superfamily. Similar to chymotrypsin, FIX is activated by limited proteolysis. This creates a novel N-terminus that can interact with the protease domain to form several important catalytic regions such as the oxyanion hole and the S1 substrate binding pocket.[4] However, in contrast to chymotrypsin, the mere creation of a novel N-terminus proves insufficient for FIXa to reach full catalytic activity. Despite activation, FIXa displays low intrinsic activity.[5,6] To acquire physiologically relevant activity, the FIXa catalytic domain requires additional interactions with Ca^{2+}, Na^{+} and with coagulation factor VIII (FVIII). Assembly with activated FVIII (FVIIIa) is accompanied by a rate enhancement that is critical for full propagation of the coagulation cascade. Dysfunction or deficiency of FVIII is associated with the bleeding disorder *Haemophilia A*.[7]

The mechanism by which cofactor interaction drives the catalytic activity remains largely unknown and one hypothesis is that cofactor interaction affects the inherent flexibility and/or conformation of the protease domain surface loops.[8,9] Within the chymotrypsin fold of the protease domain, eight unique surface loops protrude from the dual antiparallel β-barrel core towards the surface.[10] Among the coagulation factors, these loops appear variable in composition and length, as they are believed to invoke substrate recognition by formation of specificity pockets, including the S1-pocket. Interaction of the protease domain with different ligands can stabilize the protease domain and its surface loops. Such a mechanism has previously been described in thrombin.[11]

In particular the 220-loop (chymotrypsin numbering is used throughout this paper), also known as the Na^{+}-binding loop, appears to play a key role in activated protein C (APC), FXa and thrombin.[12-14] The 220-loop is centered between the N-terminus insertion site, the S1-pocket and the cofactor binding helix.[10,14-17] The relevance of the 220-loop is illustrated by mutations that affect Na^{+}-binding, S1-pocket formation, N-terminus insertion as well as protease activity.[18-22] Whether the 220-loop also modulates these properties of the FIXa protease domain remains to be

elucidated. We previously investigated the reactivity of lysine residues in the 220-loop of FIX as a function of proteolytic activation (Chapter 5). In contrast to prothrombin and FX, FIX activation did not lead to differential labelling of the 220-loop in a footprinting approach employing primary amine-directed Tandem-Mass-Tags (TMTs). However, changes in the 220-loop did occur upon irreversible binding of the peptide substrate H-Glu-Gly-Arg-chloromethylketone (EGRck) to the FIXa active site.

The close proximity of basic lysine residues and acidic glutamic acids in the 220-loop of FIX creates the possibility that an electrostatic interplay between these residues may contribute to maintaining FIXa activity (Fig. 1A). In the present paper, this hypothesis is challenged by site-directed mutagenesis and subsequent characterization of the FIX mutants. Structural implications were assessed by footprinting techniques such as TMT-labelling and hydrogen-deuterium exchange (HDX).

RESULTS

Recombinant FIXa mutants

Inspection of available FIXa crystal structures reveals short distances (<3 Å) between the basic lysine residues (Lys222 and Lys224) of the 220-loop and nearby acidic glutamic acids (Glu217, Glu219 and Glu186) (Fig. 1A). Therefore, crystal structures predict electrostatic interactions within the 220-loop. Whether or not such 220-loop interactions exist in an aqueous environment remains unclear. Nevertheless, the *Haemophilia B* database discloses that, in this region, Lys224, Glu217 and Glu219 are essential residues for FIXa function.[3] Naturally occurring mutations of Lys224 or Glu217 to an uncharged or counter-charged residue primarily lead to a moderate or severe defect (Fig. 1B).

To study the function of these residues in the FIXa 220-loop we constructed recombinant FIXa variants with alanine substitutions. These included the mutants E217A, E219A, K224A and K222A+K224A. Recombinant FIX variants were cultured in HEK293 cells and immuno-purified using an antibody directed towards the γ-carboxylated N-terminal Gla domain.[23] Following activation and subsequent purification, SDS/PAGE displayed fully activated recombinant FIXa variants (>90%) with only some intermediate FIXα left (Fig. 1C). Quantification of active sites by titration with antithrombin-III inferred an enzyme concentration similar to the value derived from the total protein concentration (Fig. 1D).

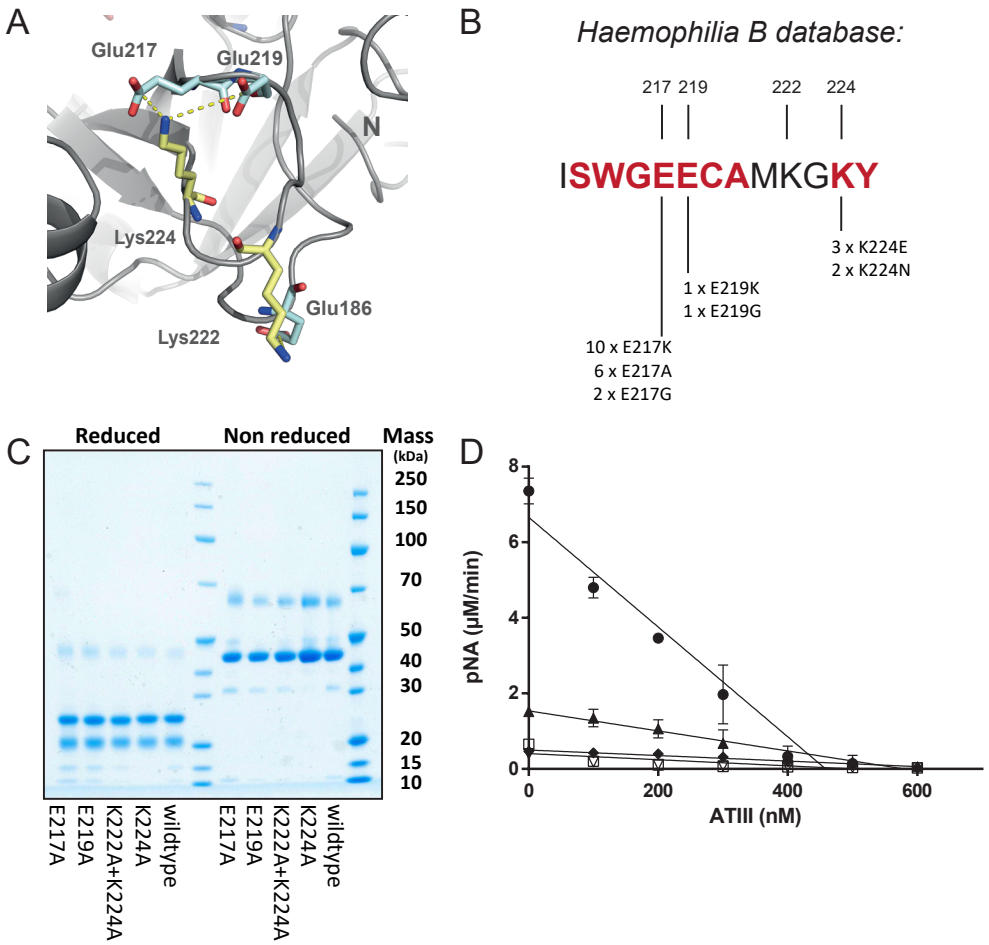


Figure 1: **220-loop variants of FIXa. (A)** Orientation of the 220-loop is displayed as derived from protein crystallography on the FIXa protease domain.[15] The charged residues Lys222, Lys224, Glu217, Glu219 and Glu186 are indicated in stick representation. Potential electrostatic interactions in this structure are indicated by a dashed, yellow line. The protease domain N-terminus is indicated (N). **(B)** A schematic representation of 220-loop mutations is given as reported in the *Haemophilia B* database.[3] Sites of various naturally occurring mutations are indicated in red, substitutions of Lys224, Glu217 and Glu219 have been described to cause bleeding classified between mild and severe. **(C)** To assess activation and purity of the FIXa variants, samples (2 µg per lane) were run on a 4-12% SDS/PAGE gel either under reducing (10 mM dithiothreitol, 95 °C for 5 minutes) or non-reducing conditions. Proteins and marker were visualized using Coomassie Brilliant Blue staining. Molecular masses of the marker are indicated on the right. **(D)** Active site determination was performed by titration with antithrombin-III in the presence of 400-600 nM of wild-type FIXa (closed circles), FIXa-E219A (closed triangles), FIXa-E217A (closed diamonds), FIXa-K224A (open triangles) and FIXa-K222A+K224A (open squares) in the presence of 2 mg/mL unfractionated heparin. Residual activity of free FIXa was quantified in 25 mM Tris, 50 mM NaCl, 5 mM $CaCl_2$ and 33% EG (pH 7.4) employing CBS 31.39 with release of pNA as described in the Experimental procedures.

Amidolytic activity of FIXa variants

Activity of variants was quantified by their ability to hydrolyse the chromogenic substrate CH_3SO_2-(D)-CHG-Gly-Arg-pNa (Pefachrome FIXa). This was assessed at physiological relevant Na^+ and Ca^{2+} concentrations. The activity towards the substrate was reduced for all mutants (Fig. 2A). Calculation of the kinetic parameters showed that the K_m for the mutants were similar to wild-type (Fig. 2A, Table I). However, a 5- to 10-fold reduction was observed in k_{cat} (Fig. 2A, Table I). This was also illustrated in terms of catalytic efficiency (k_{cat}/K_m), which showed a similar 5 to 10-fold reduction. Among the mutants, E219A and K222A+K224A appeared to be slightly more active than K224 and E217A (Table I). Remarkably, mutation of Lys222 in addition to Lys224 did not reduce activity further. In fact, a slightly higher k_{cat} was observed. These data indicate that substitution of the 220-loop lysine and glutamic acid residues greatly

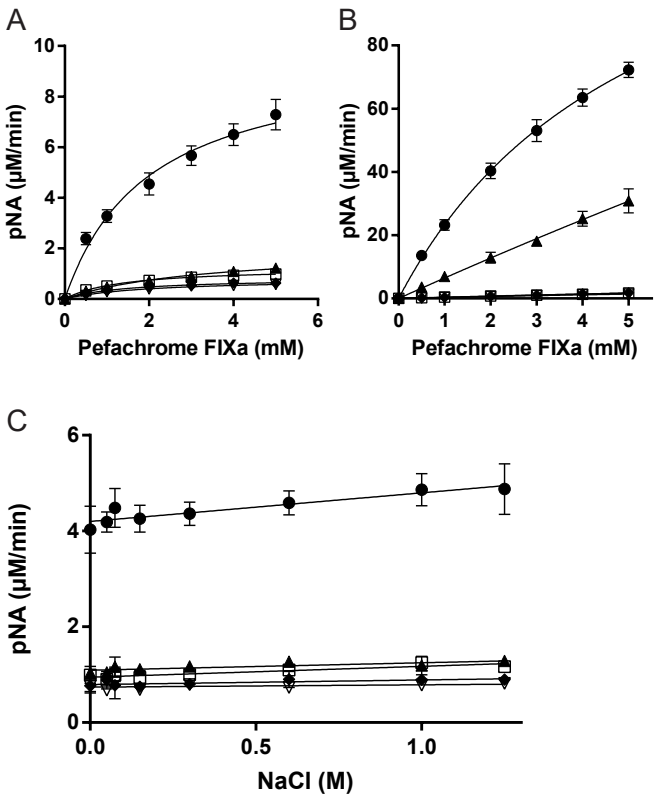


Figure 2: **Amidolytic activity of the FIXa 220-loop variants. (A)** Activity of 150 nM wild-type FIXa (closed circles), FIXa-E219A (closed triangles), FIXa-E217A (closed diamonds), FIXa-K224A (open triangles) and FIXa-K222A+K224A (open squares) towards the synthetic substrate CH_3SO_2-(D)-CHG-Gly-Arg-pNa was assessed in 50 mM Tris, 100 mM NaCl, 5 mM $CaCl_2$ and 0.2% Human Serum Albumin (pH 7.4) at 37 °C. **(B)** The same experiment was performed in the presence of 33% EG. **(C)** Amidolytic activity of FIXa variants upon titration of NaCl was assessed using of 2.5 mM of CH_3SO_2-(D)-CHG-Gly-Arg-pNa. The ionic strength was kept constant at 1.5 M with RbCl as compensatory ions. Cleavage of CH_3SO_2-(D)-CHG-Gly-Arg-pNa with concomitant pNA release was quantified as described in Experimental Procedures. Data represent the mean ± S.D. of at least three experiments.

Table I **Amidolytic activity of FIXa varaints.** The table displays the kinetic constants of the FIXa variants for hydrolysis of CH_3SO_2-(D)-CHG-Gly-Arg-pNa in the presence of Na^+ and Ca^{2+}. Constants were obtained from fitting the data of Fig. 2A and 2B using the Michaelis-Menten equation.

	Amidolytic activity		
	K_m (mM)	k_{cat} (min^{-1})	k_{cat} / K_m ($min^{-1}\ mM^{-1}$)
wild-type	2.0 ± 0.3	66.7 ± 4	33.4
E219A	3.1 ± 0.5	13.3 ± 1	4.3
E217A	1.6 ± 0.4	5.3 ± 0.6	3.3
K224A	2.0 ± 0.7	5.3 ± 0.7	2.6
K222A&K224A	1.3 ± 0.2	8.7 ± 0.5	6.7
wild-type + EG	5.4 ± 0.5	1000 ± 60	185

reduces the catalytic activity of FIXa.

In previous studies, ethylene glycol (EG) was used to potentiate the otherwise low FIXa amidolytic activity.[24,25] Crystallographic data of FIXa indicates that EG directly interacts with the 220-loop.[25] Therefore, we assessed amidolytic activity also in the presence of 33% EG (Fig. 2B). Wild-type FIXa activity showed a 5-fold enhancement in catalytic efficiency upon addition of EG (Fig. 2B, Table I). However, activity of most of the mutants remained low. E219A proved to be an exception in that its activity was at least 15-fold higher than of the other mutants (Fig. 2B). In presence of EG, the mutants displayed kinetics with an apparent K_m far exceeding the highest substrate concentrations tested. Therefore, the kinetic parameters could not be estimated.

The homologous 220-loop of thrombin, FXa and APC exhibit a strong dependency on Na^+ to maintain optimal protease activity.[12,13] Also for FIXa, Na^+ appears to interact and to some extent potentiate its activity.[26] To examine whether Na^+ is capable to recover activity of the FIXa mutants, we performed a Na^+-titration at a fixed concentration of the substrate in the presence of $CaCl_2$. Under these conditions, there was no appreciable effect of Na^+ on wild-type FIXa activity nor on the mutants (Fig. 2C).

Activity towards FX in absence and presence of FVIIIa

To further characterize the FIXa variants, their activity towards the natural substrate FX was examined in the presence of negatively charged phospholipid vesicles containing 50% phosphatidylserine (PS) and 50% phosphatidylcholine (PC). With increasing amounts of FX, wild-type FIXa displayed concentration-dependent increase in FXa formation (Fig. 3A). In contrast, the FIXa mutants showed a major

reduction in FX activation. The apparent kinetic parameters were calculated in terms of Michaelis-Menten kinetics (Table II). Wild-type FIXa exhibited a K_m of approximately 600 nM with a k_{cat} of around 121 × 10^{-3} min^{-1}. The impaired activity of the mutants was predominantly reflected by a reduced k_{cat} (Table II). Compared to wild-type FIXa, E219A showed a 12-fold lower k_{cat}, but proved to be the only mutant with appreciable activity (Fig. 3A). The other mutants, E217A, K224A and K222A+K224A, exhibited negligible activation of FX with k_{cat} being in the range of 0.1-1.5 min^{-1} (Fig. 3A, Table II).

Table II **Activity of FIXa varaints towards FX.** The table summarizes the kinetic constants of the FIXa variants for proteolysis of FX in the presence or absence of FVIII. Constants were obtained from fitting the data of Fig. 3 using the Michaelis-Menten equation.

	FX activation (in absence of FVIIIa)		FX activation (in presence of FVIIIa)		Stimulation factor
	$K_{m,app}$ (nM)	$k_{cat,app}$ × 10^{-3} (min^{-1})	$K_{m,app}$ (nM)	$k_{cat,app}$ (min^{-1})	
wild-type	600 ± 95	121 ± 8	35 ± 7	28 ± 2	230
E219A	550 ± 140	10 ± 1	nd[a]	4.2 ± 1.2	400
E217A	280 ± 100	1.2 ± 0.2	nd[a]	0.4 ± 0.1	333
K224A	322 ± 84	0.7 ± 0.1	nd[a]	0.9 ± 0.1	1300
K222A&K224A	131 ± 100	0.3 ± 0.1	nd[a]	1.5 ± 0.3	5000

[a] The $K_{m,app}$ could not be determined (nd) due to substrate inhibition.

Enhancement of FIXa activity following complex assembly with the cofactor was determined by a titration of FX in the presence of phospholipid vesicles and 0.35 nM FVIIIa. In the presence of FVIIIa, wild-type FIXa exhibited typical enhancement of its activity (Fig. 3B). The calculated kinetic constants showed a 20-fold drop in K_m and an increase in apparent k_{cat} by a factor of 230 (Table II). The apparent k_{cat} of the mutants increased by a factor of 333 or more (Table II). Despite this enhancement, FVIIIa did not recover the defect in activity and FX activation remained low compared to wild-type FIXa (Fig. 3B). In addition, upon increasing the FX concentration, substrate inhibition became apparent for the mutants carrying substitutions in the 220-loop. This substrate inhibition prohibited any estimation of K_m for the FIXa variants. Substrate inhibition is a phenomenon that had been previously reported for FIXa variants with disrupted molecular interactions.[23] Despite the lack of K_m estimates, the data in Table II show that the FIXa variants did respond to FVIIIa, however that activity remained far below the one of normal FIXa. This demonstrates that substitutions in the 220-loop are associated with reduced activity, both in the absence and presence of FVIII (Fig. 3).

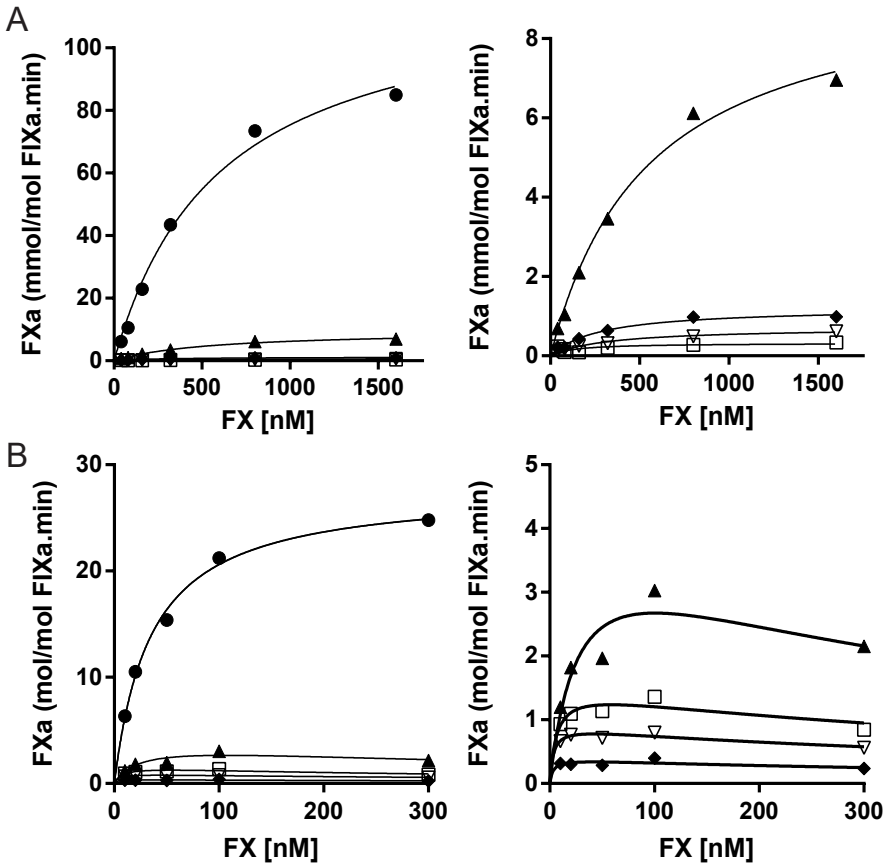


Figure 3: **Activation of FX by FIXa variants**. FX was converted to FXa by wild-type FIXa (closed circles), FIXa-E219A (closed triangles), FIXa-E217A (closed diamonds), FIXa-K224A (open triangles) and FIXa-K222A+K224A (open squares). FX was activated in presence of 50%/50% PS/PC vesicles by either **(A)** FIXa (30 nM) in absence of FVIIIa or **(B)** FIXa (0.3 nM) in presence of FVIIIa (0.35 nM). Incubations were performed at 37 °C in siliconized glass tubes containing reaction buffer (50 mM Tris, 150 mM NaCl, 10 mM $CaCl_2$ and 0.2% bovine serum albumin, pH 7.4). Separate graphs with close-ups of the data for the FIXa mutants are given on the right side. Data represent mean values of three independent experiments. Detailed description of the activations can be found in the Experimental procedures.

Chemical footprinting of protease domain N-terminus and 220-loop lysine residues

The impaired activity observed in the kinetic studies (Fig. 2 and Fig. 3) suggests a defective catalytic machinery of the FIXa mutants. To assess N-terminus insertion we studied the reactivity of Val16 in the protease domain by primary amine labelling using isobaric tandem-mass-tags (TMTs). Footprinting by TMT-labelling probes the

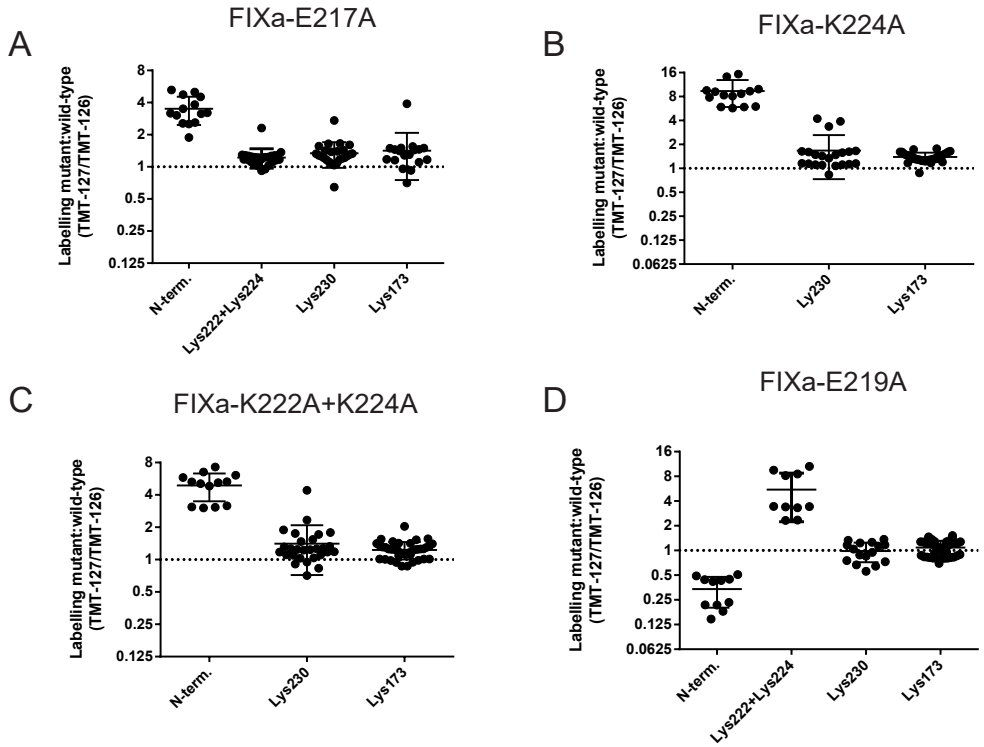


Figure 4: **TMT-labelling of the protease N-terminus and lysine residues.** The primary amine labelling is shown for **(A)** FIXa-E217A, **(B)** FIXa-K224A, **(C)** FIXa-K222A+K224A and **(D)** FIXa-E219A. Wild-type FIXa (0.5 μM) and mutant FIXa (0.5 μM) were labelled in buffer containing 20 mM HEPES, 150 mM NaCl and 5 mM $CaCl_2$ by respectively 2.5 mM of TMT-126 and TMT-127 at 25 °C for 7.5 minutes in a reaction volume of 80 μL. The reaction was quenched by incubation with 0.75% hydroxylamine for 15 min. Further quantification of labelling by mass spectrometry is described in the Experimental procedures. A pairwise comparison of labelling between mutant FIXa and wild-type FIXa is expressed in the ratio of measured TMT-labels (TMT-127/TMT-126). Individual measurements of peptides covering the N-terminus or different lysine residues are represented by black dots from triplicate experiments.

accessibility of primary amines such as the N-terminus and lysine residue ε-amines by their availability for chemical modification.[27,28] Following chemical modification, the extent of labelling was assessed using mass spectrometry. This allowed pairwise comparison between wild-type FIXa and each of the mutants, with respect to N-terminus and lysine residues labelling, using two types of TMTs (TMT-126 for wild-type and TMT-127 for mutants).

The protease domain N-terminus in K224A and K222A+K224A displayed a 4 to 8-fold higher incorporation of TMTs compared to wild-type FIXa (Fig. 4B and 4C). This indicates that the N-termini of these mutants were more solvent-exposed. Labelling of other lysine residues indicated an overall equal incorporation of TMTs.

This implies an identical label intensity in wild-type FIXa and the mutants. This is apparent from the TMT-127/TMT-126 ratio being ~1, as shown for Lys173 and Lys230. There was no information on the lysine residues of the 220-loop because these were substituted. However, increase in labelling of the protease domain N-terminus, may reflect that these lysine-to-alanine substitutions confer to FIXa a zymogen-like phenotype.

Similar to K224A and K224A+K222A, the N-terminus of E217A also proved to be more labelled compared to wild-type (TMT-127/TMT-126 ratio of ~4, Fig. 4A). This indicates that the N-terminus of E217A is more accessible for labelling than that of wild-type FIXa. In contrast, E219A was the only mutant in which labelling revealed a more protected N-terminus compared to wild-type FIXa with a TMT-127/TMT-126 ratio below 1 (Fig. 4D). Opposite to the other mutants, such protection might characterize a more proteinase-like conformation, albeit with lower activity than wild-type FIXa. Besides the N-terminus, also the lysine residues of the 220-loop can be analysed for TMT incorporation in the glutamic acid substitutions (E217A and E219A). Remarkably, in E219A, Lys222 and Lys224 proved to be more accessible for labelling (Fig. 4D). Labelling of Lys222 and Lys224 in E217A did not reveal a difference between mutant and wild-type, as the ratio between intensities of TMT-127 (E217A) and TMT-126 (wild-type) was close to 1 (Fig. 4A).

Hydrogen/deuterium exchange in the FIX protease domain

One way to probe change in structural dynamics of FIX/FIXa is to employ hydrogen/deuterium exchange (HDX). HDX determines the exchange rate of backbone amide hydrogens with surrounding deuterium when a protein is placed in D_2O. Exchange with deuterium is dependent on the interactions of the backbone hydrogens in the protein structure (i.e. H-bonding).[29,30] This technique was employed to assess change in backbone flexibility of the 220-loop and potentially of other surface loops. To study the effect of activation and reaction with the irreversible inhibitor EGRck, non-activated FIX, FIXa and FIXa-EGR were diluted in D_2O and incubated over various periods of time. Following peptic proteolysis and mass spectrometry, incorporation of deuterium was assessed.

Figure 5 displays the deuterium uptake of peptides covering most of the protease domain. The difference in deuterium uptake is given for FIX and FIXa compared to FIXa-EGR by the different coloured peptide bars (Fig. 5). Upon EGR incorporation, multiple regions were identified with a reduction in deuteration ranging from 0.5 Da to more than 2 Da (Fig. 5). Only peptides from the 90-loop region displayed a small increase in incorporation of deuterium (~0.5 Da) for the early time points. Remarkably, no appreciable difference in HDX was observed between FIX and FIXa. Changes in deuterium uptake were observed only for FIXa-EGR. This can,

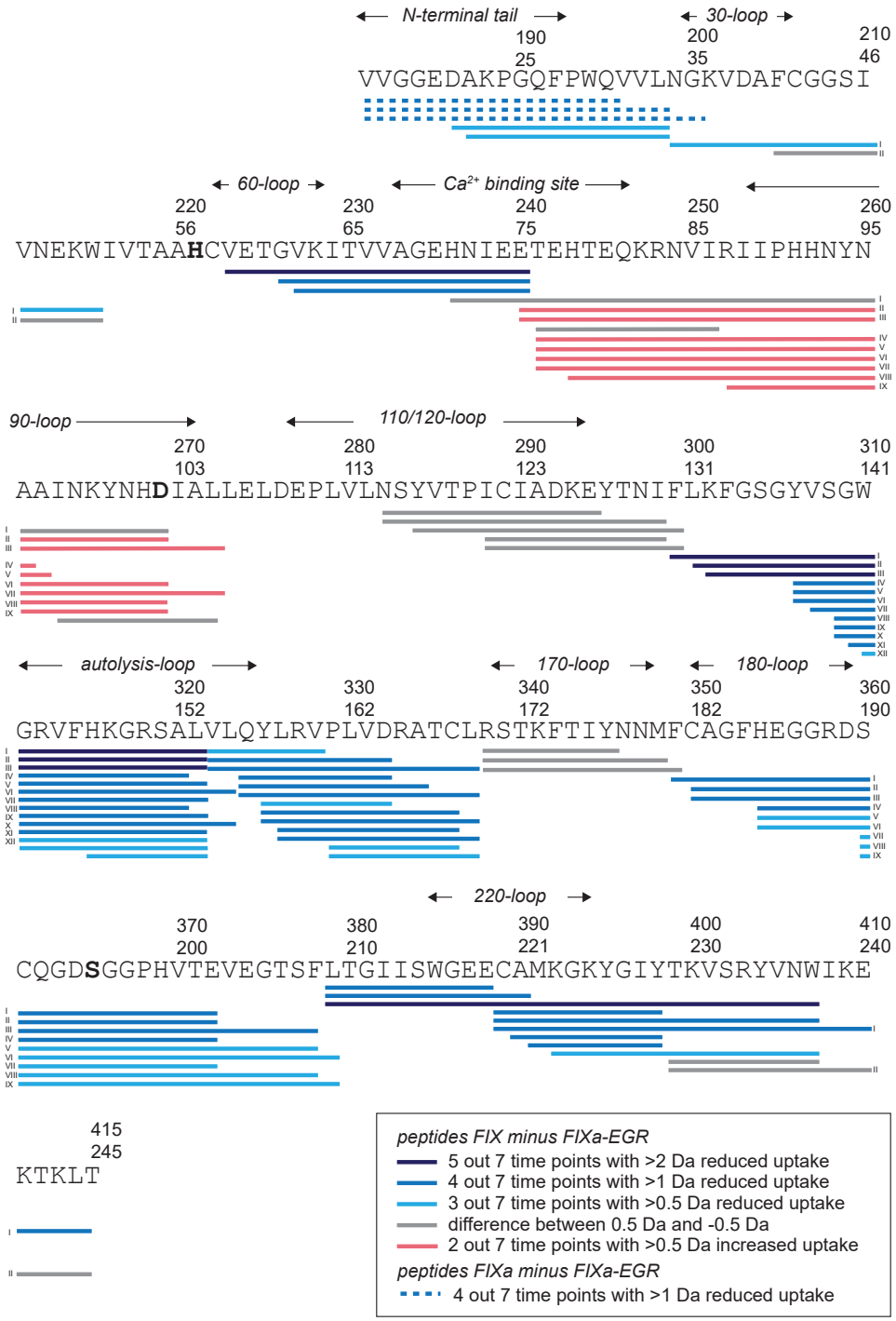


Figure 5 *(figure legend on next page)*

Figure 5: **HDX rates of peptides of the FIX protease domain.** Approximately 36 µM of FIX, FIXa and FIXa-EGR in buffer containing 200 mM HEPES, 1500 mM NaCl and 50 mM $CaCl_2$ (pH 7.1) was ten times diluted in D_2O. Incubation with D_2O at 24 °C was maintained for 10, 50, 100, 500, 10,000 and 50,000 seconds. Quenching of HDX and mass spectrometry analysis of the deuterium uptake is described in the Experimental procedures. Peptides are drawn as lines and colored according to their deuterium uptake over time. Categorization of deuterium uptake is outlined in the box in the lower right corner. Both chymotrypsin numbering (top) and FIX numbering (bottom) are given.

for instance, be seen in the individual uptake time courses for peptides acquired from the 90-region and areas with a strong reduction (Fig. 6). The areas with strong reduction in uptake are situated predominantly on side where the point of entry of the N-terminus is located (Fig. 6). These include the autolysis loop, the 180-loop, the Ca^{2+}-binding site and the 220-loop. For obvious reasons, peptides starting from the N-terminus (Val16) can be evaluated for FIXa and FIXa-EGR, but not for the uncleaved FIX zymogen. Those peptides exhibited a 1 to 2 Da reduction in deuterium uptake (Fig. 5 and 6).

When focusing on the 220-loop, FIXa-EGR displayed a reduction in uptake of about 1 Da in peptides spanning from Leu209 to Glu219. Peptides covering Cys220 to Tyr228 displayed a 3 to 4 Da decrease in uptake (Fig. 6). Reduction in deuteration was confined to an amino acid stretch ranging from Leu209 to Tyr228, whereas a peptide downstream Tyr228 (Thr229 to Tryp237) lacks the reduction in deuterium uptake (Fig. 6). Also, the changes in uptake become manifest before the first time point (< 10s) demonstrating that the incorporation of EGR affects rapidly exchanging hydrogens.[29] Thereafter, the difference in deuterium uptake remains constant throughout the incubation periods, without substantial difference in kinetics. Such consistency in uptake kinetics indicates that the rapidly exchanging amide hydrogens of FIX and FIXa, stay protected in FIXa-EGR for at least the length of the experiment (> 10.000s). With exchange rates as a measure for flexibility of the protein backbone, this persistent reduction indicates that binding of EGR, filling the S1-pocket, strongly stabilizes the 220-loop.

DISCUSSION

The aim of this study was to unravel the role of the 220-loop in the catalytic mechanism of FIXa. This particular surface loop is known to be important for catalytic activity of various serine proteases of the coagulation system. Here we made use of mass spectrometry-based footprinting techniques and mutagenesis with specific focus on the 220-loop Lys222, Lys224, Glu219 and Glu217. To determine the contribution of charged residues in the 220-loop, they were substituted to alanine residues in recombinant FIXa: E219A, E217A, K224A and K222A+K224A. In these proteins, the cleavage of a tripeptide substrate or of the natural macromolecular substrate FX appeared dysfunctional (Fig. 2A and 3A). This was to be expected since naturally occurring mutations of these residues have been reported to impair coagulation.[3] Apparently, substitutions other than replacements by alanine (see Fig. 1B) have similar effect on the proteolytic activity.

Na^+ has been reported to interact with the 220-loop and is considered to enhance activity by stabilization the 220-loop in various serine proteases.[12-14] As for FIXa, the effect of Na^+ appears insignificant in the presence of Ca^{2+} at physiological concentration.[31] In agreement with this notion, our experiments show limited effect of Na^+ on FIXa activity (Fig. 2C). More notably, Na^+ did not alleviate the defect in activity of the mutants. In the crystal structure of Na^+-bound FXa, Lys224 directly coordinates Na^+ by its backbone carbon oxygen. Nevertheless, high concentrations of Na^+ did not restore the defect introduced by K224A. Likewise, E217A and E219A did not show a response to Na^+ either. Perhaps, the presence of Lys224, Glu217 and Glu219 is required for Na^+-mediated rate enhancement. Nevertheless, considering the minor response to Na^+ in presence of Ca^{2+}, it is more likely that the mutant defect does not relate to Na^+-binding (Fig. 3C).

Similar to Na^+, EG has been reported to directly interact with the 220-loop.[25] Binding of EG is believed to stabilize the 220-loop with concomitant enhancement of the catalytic efficiency. In the EG-binding crystal published by Zögg *et al*. more structural insight is given.[25] There, the amide nitrogen of Glu217 interacts with one of the two oxygen atoms of EG. Strikingly, in our study, sidechain substitution of Glu217 or Lys224 (whether or not combined with Lys222) abolishes the rate enhancing effect of EG (Fig. 2B). In contrast, enhancement by EG seems not to involve the Glu219 side chain. Regardless of the Glu219 sidechain substitution (in E219A), EG was able to partially restore the catalytic defect (Fig. 2B). Such repair by EG suggests that the role for the charged sidechain of Glu219 in the EG rate enhancement is limited. On the contrary, the inability of EG to stimulate mutants E217A and K224A suggests a role for Glu217 and Lys224. Although structural information suggests an EG interaction with the 217 backbone, the sidechains of both Glu217 and Lys224 might be important

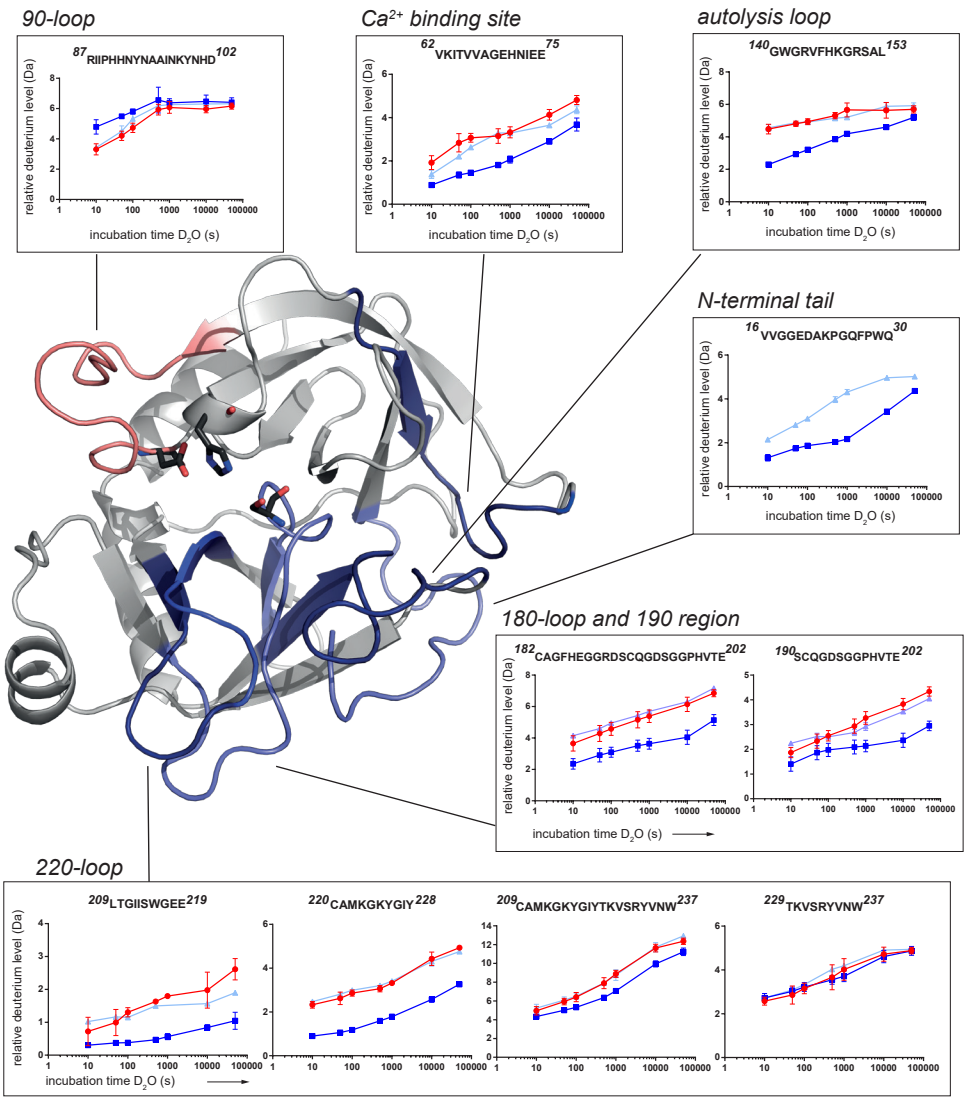


Figure 6: **Deuterium uptake and position of peptides covering surface loops and the N-terminal tail.** For a selection of peptides, the uptake of deuterium over time is plotted of FIX (solid red line), FIXa (light blue solid line) and FIXa-EGR (solid blue line). Peptides that overlap or cover the same region are combined and delineated. Chymotrypsin numbering is used to indicate the peptide boundaries and protein regions. Sites covered by the peptides are indicated on the FIXa protease domain.[15] Coloring of the crystal structure illustrates reduction (blue) or increase (light red) in deuterium uptake following binding of EGR. The catalytic triad residues His57, Asp102 and Ser195 are given in black stick representation.

in enabling the EG effect. There is no obvious relation between Lys224 and Glu217 to facilitate EG enhancement except for the close proximity of their sidechains (Fig. 1A). This raises the possibility that EG stimulation requires an electrostatic interaction between the Lys224 and Glu217 sidechains. Therefore, our EG experiments appear compatible with the concept of ion-pairing between Glu217 and Lys224 in FIXa. A functional role for a similar salt-bridge interaction has also been suggested for the homologous serine proteases thrombin and FXa.[15,18-20,22]

In an attempt to clarify the defective catalysis of mutant FIXa, TMT-labelling provided a possible explanation for E217A, K224A and K222A+K224A. Replacement of Glu217 or Lys224(+Lys222) results in an impaired insertion of the N-terminus (Fig. 4). The role of Glu219 appears more ambiguous, because E219A was the only mutant with an enhanced insertion of the N-terminus (Fig. 4D). The additional glutamic acid positioned closely to Lys224 makes FIX unique among serine proteases (Fig. 1A). Previous research supports a marked role for position 219 in the proteolytic activity of FIXa and FXa.[32,33] Surprisingly, our TMT-labelling experiment demonstrates that elimination of the charged Glu219 (E219A) sidechain drives insertion of the N-terminus (Fig. 4D). This suggests that the Glu219 sidechain, in fact, has a disadvantageous role for catalytic activity of FIXa. A mutagenesis study addressing position 219 led Hartmann *et al*. to propose that removal of the Glu219 side chain opens the S1-pocket.[34] However, in agreement with our data, mutation of Glu219 alone did not enhance FIXa catalysis and additional stabilizing mutations in the protease domain (140 and 90-loop) proved necessary to increase activity.[33,34] Possibly, in our E219A variant the active site is more open because of the substitution to alanine. It seems of interest that in E219A , TMT-labelling data shows an increased reactivity of the 220-loop lysine residues (Fig. 4D). Perhaps this indicates that the protease domain, and in particular the 220-loop therein, are more unstable in E219A compared to wild-type FIXa.

Use of HDX complements covalent TMT footprinting in that it allows to virtually visualize the entire protein and is able to provide information on the backbone hydrogen interactions.[29,30] When focusing on the 220-loop, a major reduction in hydrogen exchange can be seen upon filling of the active site by EGR (Fig. 6). The slower exchange upon substrate binding might be partially due to H-bonding between the Glu219 carbonyl group and the ArgP1 amine of EGR in the S1-pocket.[15] This projected interaction in FIXa-EGR together with the reduction in HDX suggests that the 216-220 segment stabilizes upon substrate-binding. The enhancement of activity by stabilizing surface loops (including the 220-loop) has been observed for other serine proteases in the coagulation system.[8,14,24,35-38] In our study, HDX data provided evidence that binding of a substrate analogue is accompanied by rigidification of the FIXa 220-loop.

In addition to the 220-loop, also the 180-loop, 140-loop (autolysis loop) and the N-terminal tail showed lower exchange rates upon EGR incorporation (Fig. 5 and Fig. 6). These surface loops together comprise the so-called "activation domain" because of their position close the N-terminus insertion site and observed restructuring due to activation.[39,40] Remarkably, differences in HDX could not be observed between FIX and activated FIX, (Fig. 6). A difference in exchange rates became only apparent after inactivation by EGRck. An EGR-induced reduction in HDX was also visible for other parts beyond the activation domain, including the Ca^{2+}-binding site and the 60-loop (Fig. 6). The only exception is the 90-loop, which displays an increase in HDX (Fig. 6). Previous studies have shown that stabilization of the 90-loop favours FIXa activity.[15,25,32,33] Our HDX data might corroborate the view that the 90-loop is more dynamic in FIXa-EGR. However, the difference is limited in comparison with other sites and is only visible in the early timepoints. In general, the protease domain displays reduction in exchange rates that are only noticeable in presence of the substrate analogue EGR.

Reduction in HDX might be a consequence of fewer fluctuations in the protein structure and stabilization of previously flexible or unstructured parts.[30] In that sense, lack of difference between the protease domains of FIXa and FIX might demonstrate unaltered dynamics in conformation following activation (Fig. 5 and 6). Interestingly, when HDX was performed on FVII and FVIIa the same observation was reported.[35] Although different sites were implicated (e.g. 170-loop), the mechanism for rate enhancement might be similar. The necessary structural changes for FIXa catalytic activity are not conceived by activation alone, and become manifest following substrate-binding. Such substrate-assisted changes may play a predominant role in the allosteric stabilization of the protease domain of FIXa, in analogy with FVIIa.[35]

EXPERIMENTAL PROCEDURES

Materials

Precast SDS/PAGE gels were supplied by Invitrogen (Breda, the Netherlands). Brilliant Blue Coomassie was from Invitrogen (Breda, the Netherlands). Antithrombin-III and unfractionated heparin were obtained from Sigma (Zwijndrecht, the Netherlands). Chromogenic substrate CBS 31.39 was from Stago (Leiden, the Netherlands), Pefachrome FIXa was obtained from Pentapharm (Aesch, Switzerland) and S-2765 containing the thrombin inhibitor I-2581 was from Chromogenix (Milano, Italy). PageRuler Broad Range protein marker, TMT-labels, hydroxylamine and endoprotease chymotrypsin were obtained from Thermo Scientific (Breda, the Netherlands). H-Glu-Gly-Arg-chloromethylketone (EGRck) was from Bachem (Bubendorf, Switzerland). Porcine brain L-α-phosphatidylserine (PS) and chicken egg L-α-phosphatidylcholine (PC) were from Avanti Polar Lipids, Inc. (Alabaster, AL, USA).

FIX and FIXa variants

FIX variants with amino acid substitution E217A, E219A, K224A and K222A+K224A were constructed by site directed mutagenesis in a pcDNA3.1(-) vector encoding wild-type FIX.[23] Mutagenesis was performed using the Quick Change kit (Agilent Technologies, Amstelveen, the Netherlands) following the manufacturer's instruction using appropriate primers. Mutagenesis was verified by sequencing of the FIX encoding parts on the mutant plasmids. Transfection of HEK293 cells and production of recombinant FIX variants was performed essentially as described elsewhere.[23] Immunopurification using a monoclonal antibody directed against the Gla domain (CLB-FIX 11),[41] subsequent activation and quantification of recombinant FIX variants have been described elsewhere.[23,42] Human plasma-derived FIX was obtained as immunopurified concentrate (Nonafact®, Sanquin Plasma Products, Amsterdam, the Netherlands), which was further processed by hydrophobic interaction chromatography (GE Healthcare, Eindhoven, the Netherlands) and strong anion exchange by Q-sepharose (GE Healthcare, Eindhoven, the Netherlands). FIXa-EGR was prepared by incubation of FIXa (66.5 μM) with EGRck (4 mM) for 45 minutes at 37 ºC. Excess of EGRck was removed by strong anion exchange chromatography employing Q-Sepharose as outlined previously.[23]

Determination of amidolytic activity and steady state kinetics

Cleavage of synthetic substrates with concomitant release of pNA was monitored by absorbance at 405 nm using a kinetic microtiter plate reader thermostated at 37 ºC (Molecular Devices, Wokingham, United Kingdom). Absorbance was converted to molar concentration using an extinction coefficient of 9.9 mM^{-1} cm^{-}

[1]. For FXa generation, initial rates of pNA production were converted to FXa molar concentrations using a standard curve of active-site titrated purified FXa. Kinetic constants for both the amidolytic activity and FXa generation assays were calculated according to the Michaelis-Menten equation. Fitting was done in Graphpad (Prism, La Jolla, CA, USA) using non-linear regression. Substrate inhibition was fitted using an adapted Michaelis-Menten equation described in Graphpad.

FX activation by FIXa variants

Activation of FX was essentially done as described previously.[23] Briefly, prior to activation, phospholipids (50%/50% PS/PC, 50 μM) were pre-incubated in reaction buffer for 10 minutes at 37 °C. In the absence of FVIIIa, 30 nM FIXa was added to the phospholipid mixture. FX activation was then initiated by addition of FX and allowed for 1, 2, 3 and 5 minutes. In the case of FX activation in the presence of FVIIIa, FVIII (0.35 nM final concentration) was activated for 1 minute by thrombin (1 nM), prior to the addition of FIXa (0.3 nM). FXa generation was then initiated by addition of FX and allowed for 0.5, 1, 1.5 and 3 minutes. Subsamples (50 μL) were taken from the FX activation mixtures (0.25 mL total volume) and transferred to a microtiter plate containing 200 μL buffer containing 16 mM EDTA to quench the reaction. Generated FXa was quantified using S-2765/I2581 substrate as described.[23]

Mass spectrometry analysis of TMT-labelled proteins

Following TMT-labelling, wild-type and mutant FIXa samples were mixed in a 1:1 molar ratio. Reduction, alkylation and proteolytic processing into peptides were performed as described elsewhere.[27] In order to determine N-terminus labelling, peptides were analyzed using MS^3 fragmentation in an Orbitrap Fusion mass spectrometer (Thermo Scientific, Breda, the Netherlands). Peptides of labelled FIXa variants were separated by reversed phase liquid chromatography on a C18-column packed in-house with ReproSil-Pur C18-AQ, 1.9-μm resin (Dr. Maisch, Ammerbuch-Entringen, Germany) in a 20-cm 75 to 360-μm inner- to outer-diameter fused silica emitter (New Objective, Woburn, MA). Samples were loaded for 17 minutes at 250 nL/min in 95% solution A (0.5% acetic acid) and 5% solution B (0.5% acetic acid and 80% acetonitrile). Equilibration was done for 5 minutes in the same mixture (17-22 minutes), and elution by increasing solution B from 5% to 30% (22-132 minutes) and 30% to 60% (132-147 minutes), followed by a 10-minute wash to 90% and a 5-minute regeneration to 5%. Using a nano-electrospray ion source set at 2.1 kV, peptides were sprayed into an Orbitrap Fusion mass spectrometer.

Data-dependent acquisition was initiated by a full scan in the Orbitrap with 120.000 resolution power, a scan range between 400-1500 m/z, 4.0×10^5 ion count target and maximum injection time of 50 ms. The 10 most intense precursors with a

charge state of 2 to 8 were sampled for MS^2. Dynamic exclusion duration was set to 10 seconds with a 10-ppm tolerance around the selected precursor and its isotopes. Monoisotopic precursor selection was turned on. MS^2 scans were attained following collision induced dissociation (CID) at 35% collision energy and detection in the ion-trap comprising a fragment isolation window of 1.6 m/z and 60 ms maximum injection time. The 5 most intense y, b or a ions were selected for subsequent MS^3 fragmentation to quantify the TMT labels. MS^3 fragmentation was performed by higher energy collision induced dissociation (HCD) at 65% collision energy and a MS^2 isolation width of 2 m/z. The TMT reporter groups were detected in the Orbitrap with 60000 resolution power, scan range of 100-500 m/z, 1.0×10^5 ion count target and maximum injection time of 60 ms. Parallelizable time for all available injection ions was enabled.

Peaks 7.0 (Bioinformatics Solution Inc., Waterloo, Canada) was employed to identify the N-terminal peptide V(+TMT)VGGEDAK(+TMT)PGQFPW using a database containing the human proteome (uniprot code organism 9606 and keyword kw 0181) with optional TMT modification of lysine residues and the N-terminus. To avoid interference of the labelled lysine residue on position 8, the $b6^+$-ion (representing peptide V(+TMT)VGGED) was selected for further quantification in the HCD generated MS^3 spectra. TMT-127/-TMT-126 ratios were inferred from the MS^3 spectra of the $b6^+$-ion according to peak intensities of masses 126.1274 and 127.1310 in Xcalibur 2.0 (Thermo Scientific, Breda, the Netherlands). The same method of quantification was employed on control lysine residues Lys230 and Lys173 using y and b-ions comprising the respective lysine residues.

Quantification of the complete lysine residue set, including Lys222+Lys224, was done following MS^2 mass spectrometry in an Orbitrap XL mass spectrometer (Thermo Scientific) and using Proteome Discoverer software (Thermo Scientific, Breda, the Netherlands) as described previously.[27]

Hydrogen/deuterium exchange of plasma derived FIX, FIXa and FIXa-EGR

HDX rates were acquired by use of an automated LEAP robot (LEAP technologies, Morrisville, NC, USA) that handled and timed D_2O incubation, reaction quenching, proteolytic processing, liquid chromatography and injection into the mass spectrometer for all samples.[43] Briefly, protein was diluted 10-fold in D_2O and incubated for various periods of time at 24 °C. HDX was quenched by 1:1 addition of 1.25 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 2 M Urea (pH 2.5) at 4 °C. Further processing of in-line digestion into peptides and liquid chromatography was done in a cooled environment of 4 °C. Pepsin digestion was performed by passage of the samples over a Poroszyme Immobilized Pepsin Cartridge (Thermo Scientific,

Breda, the Netherlands) using isocratic flow of 0.1% formic acid with 5% acetonitrile at 100 µL/min for 5 minutes. Generated peptides were collected on a ethyl-bridged hybrid (BEH) C18 1.7 µm VanGuard pre-column (Waters, Etten-Leur, the Netherlands). Following 30 seconds of washing, the pre-column was switched in line with a Hypersil GOLD C18 analytic column (Thermo Scientific, Breda, the Netherlands). Peptides were separated using a 12 minute gradient going from 8 to 40% employing a mobile phase of 0.1% formic acid in 80% acetonitrile under 50 µL/min flow. The mobile phase was sprayed into an Orbitrap XL mass spectrometer. Electrospray was created using an ESI source applying 4.5 kV source voltage and 30, 10 and 10 arbitrary units of gas flow of respectively sheath, auxiliary and sweep gas. For the identification of non-deuterated peptides, data dependent acquisition comprised of a full scan in the Orbitrap with 30000 resolution power, scan range of 300-2000 m/z and maximum injection time of 100 ms. The 3 most intense precursor ions were subjected to MS^2 fragmentation by collision induced dissociation at 30% normalized collision energy. Dynamic exclusion duration was set to 10 seconds with a 10-ppm tolerance around the selected precursor and its isotopes. Fragments were detected in the ion-trap with a 1.0 Da isolation width, detection range of 250-2000 m/z and maximum injection time of 30 ms. Amino acid sequence and chromatography retention times of non-deuterated peptides were resolved using PEAKS database search (PEAKS 7.0, Bioinformatics Solutions Inc., Waterloo, Canada) comprising a database with several human coagulation factors and related proteins. Corresponding deuterated peptides were measured in full scan mode of the Orbitrap at 60000 resolution power. Relative deuterium uptake was determined by calculation of the centroid mass of deuterated peptides using HDExaminer (Sierra Analytics Inc., Modesto, CA, USA). Correction for back-exchange was not applied.

REFERENCES

1. Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry.* 1991 **30** 10363-10370.
2. Stassen JM, Arnout J, Deckmyn H. The hemostatic system. *Curr Med Chem.* 2004 **11** 2245-2260.
3. 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** 1329-1340.
4. Zögg T, Brandstetter H. Activation mechanisms of coagulation factor IX. *Biol Chem.* 2009 **390** 391-400.
5. Mertens K, van Wijngaarden A, Bertina RM. The role of factor VIII in the activation of human blood coagulation factor X by activated factor IX. *Thromb Haemost.* 1985 **54** 654-660.
6. Van Dieijen G, Tans G, Rosing J, Hemker HC. The role of phospholipid and factor VIIIa in the activation of bovine factor X. *J Biol Chem.* 1981 **256** 3433-3442.
7. Fay PJ. Factor VIII structure and function. *Int J Hematol.* 2006 **83** 103-108.
8. Lechtenberg BC, Freund SM, Huntington JA. An ensemble view of thrombin allostery. *Biol Chem.* 2012 **393** 889-898.
9. Vogt AD, Chakraborty P, Di Cera E. Kinetic dissection of the pre-existing conformational equilibrium in the trypsin fold. *J Biol Chem.* 2015 **290** 22435-22445.
10. Page MJ, Di Cera E. Serine peptidases: classification, structure and function. *Cell Mol Life Sci.* 2008 **65** 1220-1236.
11. Lechtenberg BC, Johnson DJ, Freund SM, Huntington JA. NMR resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc Natl Acad Sci U S A.* 2010 **107** 14087-14092.
12. Underwood MC, Zhong D, Mathur A, Heyduk T, Bajaj SP. Thermodynamic linkage between the S1 site, the Na^+ site, and the Ca^{2+} site in the protease domain of human coagulation factor Xa. Studies on catalytic efficiency and inhibitor binding. *J Biol Chem.* 2000 **275** 36876-36884.
13. Schmidt AE, Padmanabhan K, Underwood MC, Bode W, Mather T, Bajaj SP. Thermodynamic linkage between the S1 site, the Na^+ site, and the Ca^{2+} site in the protease domain of human activated protein C (APC). Sodium ion in the APC crystal structure is coordinated to four carbonyl groups from two separate loops. *J Biol Chem.* 2002 **277** 28987-28995.
14. Huntington JA. How Na^+ activates thrombin--a review of the functional and structural data. *Biol Chem.* 2008 **389** 1025-1035.

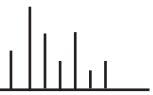
15. Hopfner KP, Lang A, Karcher A, Sichler K, Kopetzki E, Brandstetter H, Huber R, Bode W, Engh RA. Coagulation factor IXa: the relaxed conformation of Tyr99 blocks substrate binding. *Structure.* 1999 **7** 989-996.
16. Bajaj SP, Schmidt AE, Mathur A, Padmanabhan K, Zhong D, Mastri M, Fay PJ. Factor IXa:factor VIIIa interaction. Helix 330-338 of factor IXa interacts with residues 558-565 and spatially adjacent regions of the a2 subunit of factor VIIIa. *J Biol Chem.* 2001 **276** 16302-16309.
17. Kolkman JA, Lenting PJ, Mertens K. Regions 301-303 and 333-339 in the catalytic domain of blood coagulation factor IX are factor VIII-interactive sites involved in stimulation of enzyme activity. *Biochem J.* 1999 **339** 217-221.
18. Dang QD, Guinto ER, Di Cera E. Rational engineering of activity and specificity in a serine protease. *Nat Biotechnol.* 1997 **15** 146-149.
19. Gibbs CS, Coutré SE, Tsiang M, Li WX, Jain AK, Dunn KE, Law VS, Mao CT, Matsumura SY, Mejza SJ, et al. Conversion of thrombin into an anticoagulant by protein engineering. *Nature.* 1995 **378** 413-416.
20. Carter WJ, Myles T, Gibbs CS, Leung LL, Huntington JA. Crystal structure of anticoagulant thrombin variant E217K provides insights into thrombin allostery. *J Biol Chem.* 2004 **279** 26387-26394.
21. Camire RM. Prothrombinase assembly and S1 site occupation restore the catalytic activity of FXa impaired by mutation at the sodium-binding site. *J Biol Chem.* 2002 **277** 37863-37870.
22. Yang L, Manithody C, Qureshi SH, Rezaie AR. Factor Va alters the conformation of the Na^+-binding loop of factor Xa in the prothrombinase complex. *Biochemistry.* 2008 **47** 5976-5985.
23. Fribourg C, Meijer AB, Mertens K. The interface between the EGF2 domain and the protease domain in blood coagulation factor IX contributes to factor VIII binding and factor X activation. *Biochemistry.* 2006 **45** 10777-10785.
24. Sturzebecher J, Kopetzki E, Bode W, Hopfner KP. Dramatic enhancement of the catalytic activity of coagulation factor IXa by alcohols. *FEBS Lett.* 1997 **412** 295-300.
25. Zögg T, Brandstetter H. Structural basis of the cofactor- and substrate-assisted activation of human coagulation factor IXa. *Structure.* 2009 **17** 1669-1678.
26. Schmidt AE, Stewart JE, Mathur A, Krishnaswamy S, Bajaj SP. Na^+ site in blood coagulation factor IXa: effect on catalysis and factor VIIIa binding. *J Mol Biol.* 2005 **350** 78-91.
27. Bloem E, Ebberink EH, van den Biggelaar M, van der Zwaan C, Mertens K, Meijer AB. A novel chemical footprinting approach identifies critical lysine residues involved in the binding of receptor-associated protein to cluster II of LDL receptor-related protein. *Biochem J.* 2015 **468** 65-72.

28. Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem.* 2003 **75** 1895-1904.
29. Morgan CR, Engen JR. Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr Protoc Protein Sci.* 2009 Chapter 17 Unit 17.6.1-17.
30. Skinner JJ, Lim WK, Bédard S, Black BE, Englander SW. Protein dynamics viewed by hydrogen exchange. *Protein Sci.* 2012 **21** 996-1005.
31. Gopalakrishna K, Rezaie AR. The influence of sodium ion binding on factor IXa activity. *Thromb Haemost* 2006 **95** 936-941.
32. Hopfner K, Brandstetter H, Karcher A, Kopetzki E, Huber R, Engh RA, Bode W. Converting blood coagulation factor IXa into factor Xa: dramatic increase in amidolytic activity identifies important active site determinants. *EMBO Journal* 1997 **16** 6626-6635.
33. Sichler K, Kopetzki E, Huber R, Bode W, Hopfner KP, Brandstetter H. Physiological fIXa activation involves a cooperative conformational rearrangement of the 99-loop. *J Biol Chem.* 2003 **278** 4121-4126.
34. Hartmann R, Dockal M, Kammlander W, Panholzer E, Nicolaes GA, Fiedler C, Rosing J, Scheiflinger F. Factor IX mutants with enhanced catalytic activity. *J Thromb Haemost.* 2009 **7** 1656-1662.
35. Rand KD, Jørgensen TJ, Olsen OH, Persson E, Jensen ON, Stennicke HR, Andersen MD. Allosteric activation of coagulation factor VIIa visualized by hydrogen exchange. *J Biol Chem.* 2006 **281** 23018-23024.
36. Kamath P, Huntington JA, Krishnaswamy S. Ligand binding shuttles thrombin along a continuum of zymogen- and proteinase-like states. *J Biol Chem.* 2010 **285** 28651-28658.
37. Handley LD, Fuglestad B, Stearns K, Tonelli M, Fenwick RB, Markwick PR, Komives EA. NMR reveals a dynamic allosteric pathway in thrombin. *Sci Rep.* 2017 **7** 39575.
38. Croy CH, Koeppe JR, Bergqvist S, Komives EA. Allosteric changes in solvent accessibility observed in thrombin upon active site occupation. *Biochemistry.* 2004 **43** 5246-5255.
39. Huntington JA. Slow thrombin is zymogen-like. *J Thromb Haemost.* 2009 **7** 159-164.
40. Huber R, Bode W. Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* 1978 **11** 114–122
41. Christophe OD, Lenting PJ, Cherel G, Boon-Spijker M, Lavergne JM, Boertjes R, Briquel ME, de Goede-Bolder A, Goudemand J, Gaillard S, d'Oiron R, Meyer D, Mertens K. Functional mapping of anti-factor IX inhibitors developed in patients

with severe hemophilia B. *Blood.* 2001 **98** 1416-1423.

42. Lenting PJ, ter Maat H, Clijsters PP, Donath MJ, van Mourik JA, Mertens K. Cleavage at arginine 145 in human blood coagulation factor IX converts the zymogen into a factor VIII binding enzyme. *J Biol Chem.* 1995 **270** 14884-14890.
43. Andersen MD, Faber JH. Structural characterization of both the non-proteolytic and the proteolytic activation pathways of coagulation Factor XIII studied by hydrogen-deuterium exchange mass spectrometry. *Int J Mass Spectrom.* 2011 **302** 139-148.

7



Summarizing discussion

In the present thesis we explored molecular mechanisms in the haemostatic network using novel mass spectrometry footprinting techniques. These include labelling of lysine residues by use of tandem-mass-tags (TMTs) and hydrogen deuterium exchange (HDX). These experimental strategies have contributed new insight into inter- and intramolecular interactions for a selection of proteins of the haemostatic system. In this chapter, we review the major conclusions and discuss possible directions for future study.

Footprinting explored in plasma proteins

In Chapter 2, the interaction of low-density lipoprotein receptor-related protein (LRP) with its natural antagonist receptor-associated protein (RAP) was utilized as model system to evaluate the exposure of lysine residues by isobaric tandem mass tags (TMTs). With the use of the TMTs, we were able to compare lysine exposure of RAP either in complex with LRP or without LRP. This resulted in the successful identification of two lysine residues that are known to be critical for complex formation between RAP and LRP. Surprisingly, however, we also found that two additional lysine residues, at position 191 (D2 domain) and 306 (D3 domain), contribute to LRP binding. Recently, the same sites were recovered by a different approach, involving cross-linking lysine residues of RAP and LRP.[1] RAP was of particular interest because of the characteristic way its lysine residues interact via the so-called acidic necklaces in LRP. These typical binding pockets provide a general mechanism that supports the assembly of LRP with its many other ligands, and RAP can be considered as a representative example thereof. For instance, a similar binding mechanism has been proposed for other plasma proteins that are scavenged by LRP such as factor IX (FIX) and factor VIII (FVIII).[2,3] It has recently been resolved how FVIIIa is capable of binding LRP through an interface containing surface exposed lysine residues.[4,5] Perturbation of these lysine residues interactions might eventually be useful to decrease cellular uptake of the scavenged coagulation factors.[6] A complicating factor, however, is that other lysine residues can replace the interaction with LRP following removal of critical lysine residues (Chapter 2).[7] The comparative approach with TMTs may be useful in identifying such compensatory lysine residues.

In Chapter 3, TMT footprinting was initially used to assess domain organization in FVIII chimeras carrying C-domains from factor V. Swapping of each of the two C-domains induced increased exposure of certain lysine residues, which led us to hypothesize a change in connectivity in the paired tandem structure of the C-domains, leading to an increased mobility of the C2-domain therein. Surprisingly, this apparent flexibility within the C-domain pair had limited impact on FVIII function and did not affect the interaction with von Willebrand factor (VWF) and FIXa. However, expression of FVIII/V chimeras in endothelial cells revealed altered

intracellular trafficking. It might be interesting to further unravel the role of C-domain orientation in interactions of FVIII with other partners in the haemostatic network. For instance, (zero-length) lysine cross-linking techniques may be useful to see if the probed lysine residues actually make interdomain contacts.[8] Perhaps, in a reciprocal way, cross-linking interdomain contacts and fixation of the C-domain tandem pair affects binding to VWF, FIXa or even cell surface receptors differently. Such subtle changes in C-domain mobility might be useful in designing next-generation FVIII variants with improved pharmacokinetics.

Chapter 4 focuses on a serine protease from the haemostatic network called hyaluronan-binding protein 2 (HABP2). In this chapter, we present a combinatorial approach of N-terminus footprinting, kinetic studies and computer modelling to address the putative pathogenicity of the HABP2 Marburg-1 (G221E) polymorphism. Proteolytic activation of HABP2, like in other serine proteases, generates a new N-terminus at the protease domain. Insertion thereof into the active site is generally considered as a hallmark of conversion of the inactive zymogen into the enzymatically active serine protease. Footprinting using a biotin-carrying tag revealed that substitution of Gly221 to glutamic acid impairs insertion of the N-terminus of the activated HABP2 and thus explains the impaired enzymatic activity of the Marburg-1 variant. Homology modelling suggested that N-terminus insertion in the activated HABP2 variant is sterically hindered by the side chain of the glutamic acid at position 221. This chapter illustrates the potential of footprinting in studying protease activation and, in particular, the molecular events that are associated with zymogen to enzyme transition. This theme is further explored in the following chapters.

Allosteric mechanisms in coagulation serine proteases explored by chemical footprinting

Chapter 5 describes the conversion of prothrombin to thrombin as explored by comparative footprinting employing TMTs. Major differences were observed in lysine reactivity near the C-terminus of the protease domain. Lysine residues in the exosite II, 140-, 180- and 220-loop showed an activation-dependent increase or decrease in reactivity. In combination with available structural information, one may speculate on the function of the probed lysine residues. For instance, the data for two particular lysine residues, at positions 185 and 186d, indicate that the 180-loop undergoes a dramatic change in exposure. Structural studies show that the unusually long 180-loop of prothrombin reshapes in thrombin to create a conspicuous 'water channel' towards the active site (Fig. 1).[9-11] Footprinting suggests that this might be an activation-dependent mechanism, perhaps with a direct role for the lysine residues therein. Anyway, the lysine residues prove to be useful molecular sensors to monitor such changes.

We employed the same time-resolved footprinting approach on FX and FIX to explore differences in activation between these serine proteases. Of the probed changes, lysine residues in the 220-loop of thrombin and FXa displayed a decrease in exposure. In contrast, activation of FIX did not display such alterations in its 220-loop lysine residues. The 220-loop in thrombin contributes to Na^+-binding, which is essential for its catalytic activity. This Na^+-effect underscores the relevance of the 220-loop for full catalytic activity. Interestingly, while the 220-loop of FXa displays thrombin-like changes, this does not occur in FIX activation. In this respect, FIXa seems more zymogen-like than FXa or thrombin, which would justify its particularly low enzymatic activity.

A typical structural feature of serine protease activation is folding of the N-terminus into the protease domain near the active site. Reorganization of the active site, following interaction with the N-terminus, has long time been regarded as a key element in conversion of inactive zymogens to active proteases. Among the serine proteases studied in this thesis, the available crystallographic data display an inserted N-terminus, also when structures were obtained in absence of a pseudo-substrate (Fig. 2). Therefore, structural information supports the widely accepted paradigm of a tight insertion of the N-terminus following activation.[12] Footprinting of the N-terminus, however, demonstrated a more complex view. In Chapter 5, N-terminus accessibility of wild-type FIXa, FXa and thrombin was compared to their pseudo-substrate bound forms. Prior to substrate binding, the N-terminus appeared considerably more susceptible for labelling than after incorporation of an irreversible inhibitor into the active site (Chapter 5). Thus, in contrast to the view derived from crystallographic information, N-terminus insertion is a more dynamic process that may also involve substrate binding. The complexity of this allosteric network is further underlined by our finding indicating that N-terminus insertion also involves the 220-loop, as became evident from mutagenesis in HABP2 (Chapter 4) and FIXa (Chapter 6). Recently, by means of relaxation dispersion NMR, others have also obtained evidence in favor of a dynamic thrombin N-terminus.[13] It therefore seems evident that N-terminus insertion is not the sole determinant for catalytic activity. It has been reported that, for catalysis to take place, other (structural) events can overtake N-terminus insertion.[14] This might explain why a zymogen serine protease like tissue-type plasminogen activator (t-PA) displays substantial activity.[15] Apparently, N-terminus insertion is one, but not the sole structural aspect that drives the catalytic activity of serine proteases.

Although valuable for studying N-terminus insertion and lysine exposure, our TMT- approach is limited in that it gives no information on structure elements that do not contain lysine residues. This limitation does not apply to HDX which gives information about virtually the entire protein. HDX, however, provides an analysis of a single protein in a static 'end-point' state, such as enzyme plus or minus inhibitor,

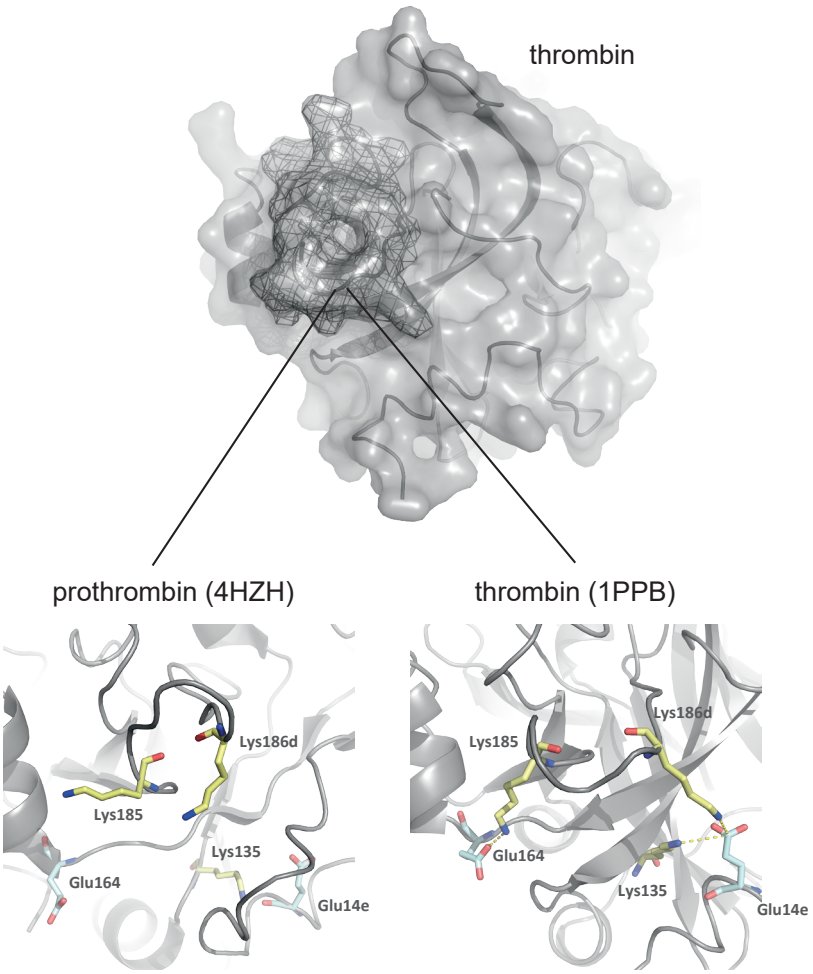


Figure 1: **Changes in 180-loop of prothrombin probed by lysine labelling.** Side view of the thrombin protease domain. The entrance of a conspicuous water channel is indicated by a dark mesh. This aperture is formed by both the 220- and 180-loop. In the crystal structure of prothrombin, lysine residues of the 180-loop (Lys185 and Lys186d) are free towards the solvent.[37] After activation, the thrombin crystal structure indicates close contacts of the same lysine residues with glutamic acids of surrounding loops.[38] This opens the possibility that, as a result of activation, the 180-loop reshapes to create the aperture. Activation dependent formation of these interactions are in line with TMT-labelling that monitors prothrombin activation (Chapter 5).

whereas our TMT-approach allows for monitoring of kinetic processes such as zymogen activation over time. Therefore, these techniques are fully complementary, and combination thereof may prove particularly beneficial. This became apparent in assessing activation of FIX in Chapter 5 and 6.

Based on TMT-labelling (Chapter 5) and HDX data (Chapter 6), solvent accessibility of the surface loops and in particular of the 220-loop seemed to be indistinguishable in both FIX and FIXa, despite release of the activation peptide, the exposure of the new N-terminus and the possible interaction with Na^+. Only upon binding of the pseudo-substrate EGR and filling of the S1-pocket, the 220-loop in FIXa displayed a reduction in backbone accessibility as detected by HDX (Chapter 6). Simultaneously, lysine residues in the 220-loop decreased in reactivity towards TMTs and the N-terminus became more protected (Chapter 5). This implies that in FIX the flexible 220-loop is stabilized by filling of the active site, but not by activation. This is reminiscent to substrate-assisted catalysis, as proposed earlier for FVIIa.[16]

Recent studies have proposed that allosteric regulation of thrombin involves thermodynamic 'ensembles' of conformational states in the transition of zymogen to protease. Binding of various ligands, including Na^+, to the allosteric binding sites then shifts the equilibrium of an ensemble of conformations towards the active configuration.[17,18] Perhaps, the low catalytic activity of FIXa is to be explained by a lack of focus towards an active state. One unique feature of FIXa is that it carries an additional glutamic acid in its putative Na^+-binding 220-loop. The presence of Glu at position 219 may produce additional zymogen-like conformations of the 220-loop and its surrounding N-terminus and S1-pocket. These additional conformations might not be present in thrombin or FXa, because these two enzymes have a small, nonpolar glycine residue at position 219. In FIXa, Glu219 might complicate the transition towards an enzyme-like configuration of the 220-loop.

In an effort to study the 220-loop in Chapter 6, we have mutated the 220-loop in such a way that the previously proposed flexible protease domain might be shifted towards certain unfavorable (sets of) conformations. For instance, substitution of Glu217 (E217A) might drive Glu219 closer to Lys224 (Fig. 3B). In this case, footprinting data indicate a less stable N-terminus insertion. Contrary, the substitution of Glu219 (E219A) might render a 219 position which is slightly removed from Lys224, but closer to the S1-pocket (Fig. 3C). Together with the footprinting data, this conformation suggests that orientation of the 219 backbone drives N-terminus insertion. Unfortunately, deduction of a linkage between the 220-loop and N-terminus insertion remains challenging based on the available structural data. Furthermore, the catalytic activity of E219A remains remarkably lower compared to the one of normal FIXa. This once more supports the notion that N-terminus insertion is not the sole requirement for catalysis.

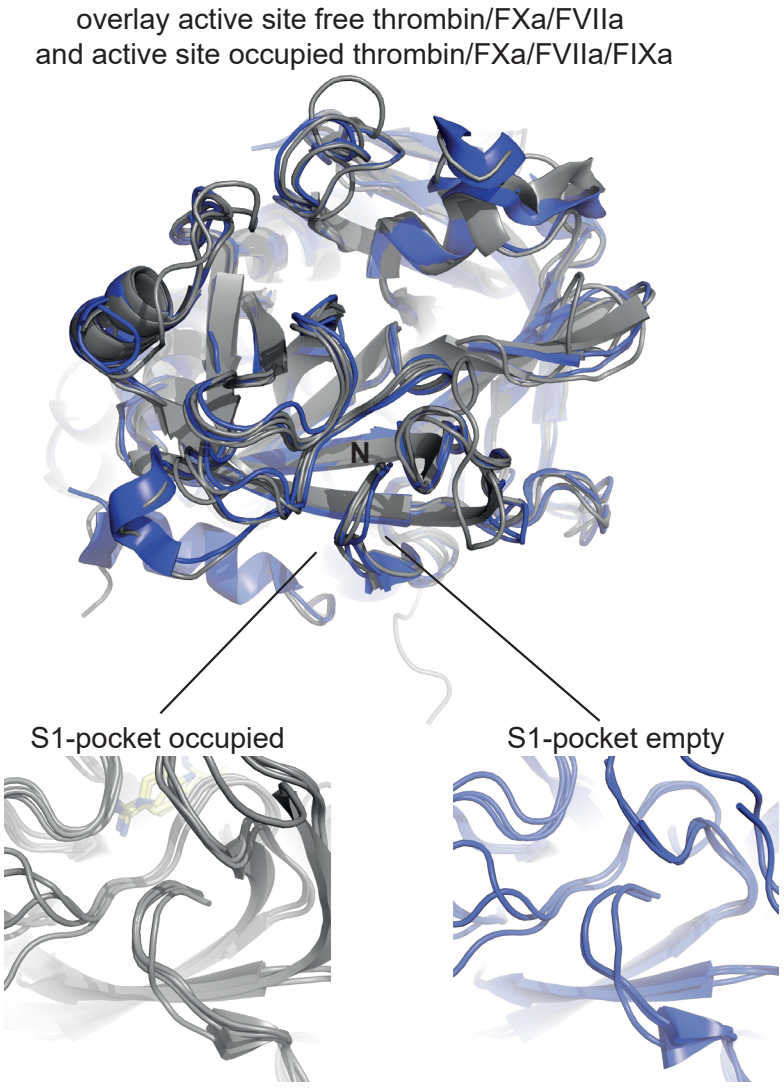


Figure 2: **N-terminus insertion observed by X-ray crystallography.** Overlay of coagulation serine proteases FIXa, FXa, FVIIa and thrombin. The FXa, FVIIa and thrombin, structures are given either with an inhibitor in the active site (indicated in grey) or an empty active site (indicated in blue). In all cases the N-terminus is folded into the protease domain. Separate close ups of the N-terminus insertion site are provided. The PDB codes used to view structures of filled active site FVIIa, FXa, FIXa and thrombin are 1KLI, 2BOK, 1RFN and 1PPB respectively.[16,38-40] PDB codes of empty active site FVIIa, FXa and thrombin are 1KLJ, 1HCG and 3U69 respectively.[16,41,42] In the case of FVIIa (1KLJ), benzamidine was initially used to crystallize the protein and subsequently soaked out.

Much remains to be learned about the enhancement of catalytic activity in plasma serine proteases. In the case of FIXa and FXa, the contribution of the Gla domain and the two EGF-like domains to the catalytic activity remains an open issue. Interestingly, several mutations in these domains affect activity of the protease domain.[19] The inferred lines of 'communication' remain to be established. Also, the effect that the cofactor has on the protease domain architecture remains elusive. There are some indications that cofactor FV and FVIII interact with an α-helix nearby the site of N-terminus insertion in FXa and FIXa.[20-24] It would be interesting to explore the effect of these cofactors on the N-terminus and the surface loops of the protease domain.

Challenges in mass spectrometry footprinting and future directions

Surface labelling by introduction of covalently bound tags has met with concerns on protein integrity.[25] The inherent change on amino acid charge and size by modification might affect protein structure on its own. Such artefactual perturbation of protein architecture could also influence subsequent labelling events and present a false outcome. Therefore, additional approaches are employed to ensure protein integrity. Techniques such as fluorescence spectroscopy, circular dichroism and activity assays have been commonly used.[25] Unfortunately, such methods generate data on a global level and are insensitive to local or subtle changes. In addition, the quantification of non-isobaric tags can be challenged by artefacts. Residues at the protein surface are more easily labelled than residues situated in the inner core. Also, differences in protein concentration could introduce variation in labelling. Here we tried to overcome these concerns by use of isobaric TMTs in a comparative way. Proteins are essentially labelled in the same way by essentially the same tags. The only difference is in the experimental variable, such as incubation time or presence of a ligand or inhibitor.

Nonetheless, a challenge in the assessment of protein complexes by TMT footprinting is their inherent dissociation and re-association kinetics. Upon transient dissociation of the protein complex, labelling of otherwise protected residues may impair the re-association. This might create a bias in the chemical footprinting. Moreover, labelling itself can drive dissociation. It is therefore important to optimize the labelling reaction for individual proteins and protein complexes. For instance, the LRP-RAP complex formation was examined by TMT footprinting for 15 minutes at 37 °C. Serine protease conformational changes appeared more sensitive and required shorter labelling times (7.5 minutes) at lower temperature (25 °C). These issues do not concern HDX because of the small and reversible nature of hydrogen exchange. However, to probe subtle changes, variation in pH, temperature and labelling time have been applied. Finetuning of such conditions might be considered for future covalent labelling studies as well.

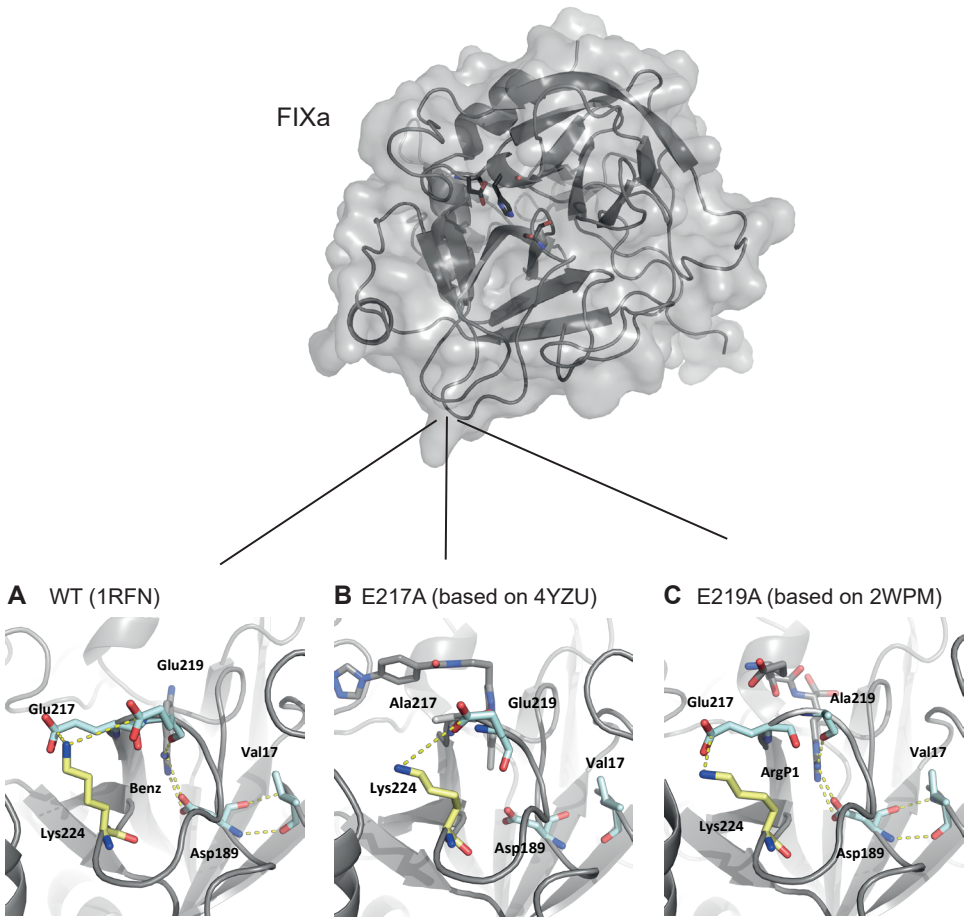


Figure 3: **Putative arrangements of the FIXa 220-loop.** Three structures with a different conformation of position 219 that might be reflected in the 220-loop variants of Chapter 6. **(A)** Close-up of the 220-loop in wild-type FIXa (PDB: 1RFN).[39] **(B)** In E217A, the substitution with alanine has been modelled using PyMol software. This structure is based on a X-ray structure of FIXa with an empty S1-pocket (PDB: 4YZU).[43] In this structure, Glu219 is positioned slightly closer towards Lys224. An interaction between Glu219 and Lys224 might compensate for the loss of the negatively charged sidechain of Glu217. **(C)** Similar to E217A, a structure of E219A was created. The original structure is based on FIXa with a filled S1-pocket (PDB: 2WPM).[44] In this case Glu219 is positioned away from Lys224 but closer to the ArgP1 of the inhibitor EGR.

The TMT and HDX footprinting approaches enabled us to unravel protein-protein interactions (LRP-RAP), subdomain orientation (FVIII/FV chimeras) and allosteric mechanisms (FIXa, FXa and thrombin). In this thesis, available structural information assisted in providing context for the probed changes. Nonetheless it remains challenging to translate that into protein function and dynamics. For instance, in many cases presented here, crystal structures indicated that probed lysine residues are in close contact with nearby glutamic acids (Fig. 1). As a consequence, lysine exposure could be interpreted as opening or closing of salt-bridge interactions. This interpretation, however, remains speculative. Many of these putative lysine-glutamic acid interactions were found at the protein surface. The contribution of such surface salt-bridges in an aqueous environment remain a matter of debate. Numerous surface salt-bridges have been identified with minimal contribution to protein stability, while some are found to stabilize proteins or even to possess a regulatory function.[26-29] To avoid the risk of overinterpretation, additional characterization of the putative salt-bridges remains necessary. For the moment, however, we believe that lysine labelling by use of TMTs has the potential to become a complementary technique next to the conventional approach of characterizing salt-bridges.

In principle, mass spectrometry footprinting can be done on proteins and protein complexes under physiological relevant conditions without restriction on size. Unfortunately, HDX has met issues in protein digestion, especially in heavily cysteine-crosslinked regions and resolution of peptides in analysis of deuterium uptake. This is due to inefficient protein digestion under cold and acidic conditions and poor peptide separation by fast liquid chromatography, in order to prevent back-exchange of deuterium to hydrogen. With the increase in sensitivity and resolving power of mass spectrometry, we might be able to study more complex systems (e.g. FIXa with FVIIIa and FX) and with smaller amounts of protein. Also, different strategies have been developed to improve digestion by electrochemical reduction, the use of novel proteases and to improve chromatography for peptide separation.[30-32] Alternatively, to surpass the digestion step, disintegration of intact proteins using different fragmentations techniques such as electron transfer dissociation (ETD) has been successfully applied, as reported for FVIIa.[33]

Besides continuous labelling, pulsed duration of HDX is commonly used to assess protein folding.[34] Therein, a denatured protein is allowed to refold under acidic conditions before being subjected to a short-lived pulse (ms to s) at high pH for HDX. A similar pulsed fashion of HDX might be useful to monitor reactions in the coagulation system. Slow activation of FIX in D_2O can be performed under inefficient labelling conditions such as a low temperature. A short duration of higher temperature then allows for efficient HDX. This could be repeated along a variation of activation times. Perhaps, such approaches may reveal transient changes in conformation or flexibility

upon FIX-to-FIXa conversion.

Even with ongoing improvements and experiments in footprinting techniques, structural and functional information remains vital to facilitate interpretation and gain understanding of the observed changes. Therefore, integration of mass spectrometry footprinting with other structural biology techniques seems rising.[35,36] As shown in this thesis footprinting in combination with functional assays and structural modelling can play a valuable role in elucidating protein interactive sites and molecular networks.

REFERENCES

1. De Nardis C, Lössl P, van den Biggelaar M, Madoori PK, Leloup N, Mertens K, Heck AJ, Gros P. Recombinant Expression of the Full-length Ectodomain of LDL Receptor-related Protein 1 (LRP1) Unravels pH-dependent Conformational Changes and the Stoichiometry of Binding with Receptor-associated Protein (RAP). *J Biol Chem.* 2017 **292** 912-924.
2. Neels JG, van Den Berg BM, Mertens K, ter Maat H, Pannekoek H, van Zonneveld AJ, Lenting PJ. Activation of factor IX zymogen results in exposure of a binding site for low-density lipoprotein receptor-related protein. *Blood.* 2000 **96** 3459-3465.
3. Bovenschen N, Herz J, Grimbergen JM, Lenting PJ, Havekes LM, Mertens K, van Vlijmen BJ. Elevated plasma factor VIII in a mouse model of low-density lipoprotein receptor-related protein deficiency. *Blood.* 2003 **101** 3933-3939
4. Young PA, Migliorini M, Strickland DK. Evidence That Factor VIII Forms a Bivalent Complex with the Low Density Lipoprotein (LDL) Receptor-related Protein 1 (LRP1): identification of cluster II on LRP1 as the major binding site. *J Biol Chem.* 2016 **291** 26035-26044.
5. Van den Biggelaar M, Madsen JJ, Faber JH, Zuurveld MG, van der Zwaan C, Olsen OH, Stennicke HR, Mertens K, Meijer AB. Factor VIII Interacts with the Endocytic Receptor Low-density Lipoprotein Receptor-related Protein 1 via an Extended Surface Comprising "Hot-Spot" Lysine Residues. *J Biol Chem.* 2015 **290** 16463-16476.
6. Meems H, van den Biggelaar M, Rondaij M, van der Zwaan C, Mertens K, Meijer AB. C1 domain residues Lys 2092 and Phe 2093 are of major importance for the endocytic uptake of coagulation factor VIII. *Int J Biochem Cell Biol.* 2011 **43** 1114-1121.
7. Prasad JM, Young PA, Strickland DK. High Affinity Binding of the Receptor-associated Protein D1D2 Domains with the Low Density Lipoprotein Receptor-related Protein (LRP1) Involves Bivalent Complex Formation: CRITICAL ROLES OF LYSINES 60 AND 191. *J Biol Chem*. 2016 **291** 18430-18439.
8. Sinz A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom Rev.* 2006 **25** 663-682.
9. Guinto ER, Caccia S, Rose T, Fütterer K, Waksman G, Di Cera E. Unexpected crucial role of residue 225 in serine proteases. *Proc Natl Acad Sci U S A.* 1999 **96** 1852-1857.
10. Pineda AO, Carrell CJ, Bush LA, Prasad S, Caccia S, Chen ZW, Mathews FS, Di Cera E. Molecular dissection of Na^+ binding to thrombin. *J Biol Chem.* 2004 **279**

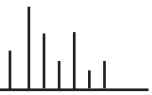
31842-31853.

11. Silva FP Jr, Antunes OA, de Alencastro RB, De Simone SG. The Na^+ binding channel of human coagulation proteases: novel insights on the structure and allosteric modulation revealed by molecular surface analysis. *Biophys Chem.* 2006 **119** 282-294.
12. Fehlhammer H, Bode W, Huber R. Crystal structure of bovine trypsinogen at 1-8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J Mol Biol.* 1977 **111** 415-438.
13. Handley LD, Fuglestad B, Stearns K, Tonelli M, Fenwick RB, Markwick PR, Komives EA. NMR reveals a dynamic allosteric pathway in thrombin. *Sci Rep.* 2017 **7** 39575.
14. Persson E, Bak H, Østergaard A, Olsen OH. Augmented intrinsic activity of Factor VIIa by replacement of residues 305, 314, 337 and 374: evidence of two unique mutational mechanisms of activity enhancement. *Biochem J.* 2004 **379** 497-503.
15. Tachias K, Madison EL. Converting tissue-type plasminogen activator into a zymogen. *J Biol Chem.* 1996 **271** 28749-28752.
16. Sichler K, Banner DW, D'Arcy A, Hopfner KP, Huber R, Bode W, Kresse GB, Kopetzki E, Brandstetter H. Crystal structures of uninhibited factor VIIa link its cofactor and substrate-assisted activation to specific interactions. *J Mol Biol.* 2002 **322** 591-603
17. Vogt AD, Pozzi N, Chen Z, Di Cera E. Essential role of conformational selection in ligand binding. *Biophys Chem.* 2014 **186** 13-21.
18. Lechtenberg BC, Freund SM, Huntington JA. An ensemble view of thrombin allostery. *Biol Chem.* 2012 **393** 889-898.
19. Fribourg C, Meijer AB, Mertens K. The interface between the EGF2 domain and the protease domain in blood coagulation factor IX contributes to factor VIII binding and factor X activation. *Biochemistry.* 2006 **45** 10777-10785.
20. Yang L, Manithody C, Qureshi SH, Rezaie AR. Factor Va alters the conformation of the Na^+-binding loop of factor Xa in the prothrombinase complex. *Biochemistry.* 2008 **47** 5976-5985.
21. Rudolph AE, Porche-Sorbet R, Miletich JP. Definition of a factor Va binding site in factor Xa. *J Biol Chem*. 2001 **276** 5123-5128.
22. Venkateswarlu D. Structural insights into the interaction of blood coagulation co-factor VIIIa with factor IXa: a computational protein-protein docking and molecular dynamics refinement study. *Biochem Biophys Res Commun.* 2014 **452** 408-414.
23. Bajaj SP, Schmidt AE, Mathur A, Padmanabhan K, Zhong D, Mastri M, Fay PJ. Factor IXa:factor VIIIa interaction. Helix 330-338 of factor IXa interacts with residues 558-565 and spatially adjacent regions of the a2 subunit of factor VIIIa. *J Biol Chem.* 2001 **276** 16302-16309.

24. Lechtenberg BC, Murray-Rust TA, Johnson DJ, Adams TE, Krishnaswamy S, Camire RM, Huntington JA. Crystal structure of the prothrombinase complex from the venom of Pseudonaja textilis. *Blood*. 2013 **122** 2777-2783.
25. Zhou Y, Vachet RW. Covalent labeling with isotopically encoded reagents for faster structural analysis of proteins by mass spectrometry. *Anal Chem*. 2013 **85** 9664-9670.
26. Makhatadze GI, Loladze VV, Ermolenko DN, Chen X, Thomas ST. Contribution of surface salt bridges to protein stability: guidelines for protein engineering. *J Mol Biol.* 2003 **327** 1135-1148.
27. Sarakatsannis JN, Duan Y. Statistical characterization of salt bridges in proteins. *Proteins.* 2005 **60** 732-739.
28. Kumar S, Nussinov R. Salt bridge stability in monomeric proteins. *J Mol Biol.* 1999 **293** 1241-1255.
29. Hinzman MW, Essex ME, Park C. Salt bridge as a gatekeeper against partial unfolding. *Protein Sci.* 2016 **25** 999-1009.
30. Mysling S, Salbo R, Ploug M, Jørgensen TJ. Electrochemical reduction of disulfide-containing proteins for hydrogen/deuterium exchange monitored by mass spectrometry. *Anal Chem.* 2014 **86** 340-345.
31. Wales TE, Fadgen KE, Eggertson MJ, Engen JR. Subzero Celsius separations in three-zone temperature controlled hydrogen deuterium exchange mass spectrometry. *J Chromatogr A.* 2017 pii: S0021-9673(17)30827-0.
32. Tsiatsiani L, Akeroyd M, Olsthoorn M, Heck AJR. Aspergillus niger Prolyl Endoprotease for Hydrogen-Deuterium Exchange Mass Spectrometry and Protein Structural Studies. *Anal Chem.* 2017 **89** 7966-7973
33. Song H, Olsen OH, Persson E, Rand KD. Sites involved in intra- and interdomain allostery associated with the activation of factor VIIa pinpointed by hydrogen-deuterium exchange and electron transfer dissociation mass spectrometry. *J Biol Chem.* 2014 **289** 35388-35396.
34. Konermann L, Simmons DA. Protein-folding kinetics and mechanisms studied by pulse-labeling and mass spectrometry. Mass Spectrom Rev. 2003 **22** 1-26.
35. Ornes S. Let the structural symphony begin. *Nature*. 2016 **536** 361-363.
36. Lössl P, van de Waterbeemd M, Heck AJ. The diverse and expanding role of mass spectrometry in structural and molecular biology. *EMBO J.* 2016 **35** 2634-2657.
37. Pozzi N, Chen Z, Gohara DW, Niu W, Heyduk T, Di Cera E. Crystal structure of prothrombin reveals conformational flexibility and mechanism of activation. J *Biol Chem.* 2013 **288** 22734-22744.
38. Bode W, Mayr I, Baumann U, Huber R, Stone SR, Hofsteenge J. The refined 1.9 Å crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment.

EMBO J. 1989 **8** 3467-3475.

39. Hopfner KP, Lang A, Karcher A, Sichler K, Kopetzki E, Brandstetter H, Huber R, Bode W, Engh RA. Coagulation factor IXa: the relaxed conformation of Tyr99 blocks substrate binding. *Structure.* 1999 **7** 989-996.
40. Schärer K, Morgenthaler M, Paulini R, Obst-Sander U, Banner DW, Schlatter D, Benz J, Stihle M, Diederich F. Quantification of cation-pi interactions in protein-ligand complexes: crystal-structure analysis of Factor Xa bound to a quaternary ammonium ion ligand. *Angew Chem Int Ed Engl.* 2005 **44** 4400-4404.
41. Padmanabhan K, Padmanabhan KP, Tulinsky A, Park CH, Bode W, Huber R, Blankenship DT, Cardin AD, Kisiel W. Structure of human des(1-45) factor Xa at 2.2 Å resolution. *J Mol Biol.* 1993 **232** 947-966.
42. Figueiredo AC, Clement CC, Zakia S, Gingold J, Philipp M, Pereira PJ. Rational design and characterization of D-Phe-Pro-D-Arg-derived direct thrombin inhibitors. *PLoS One.* 2012 **7** e34354
43. Parker DL Jr, Walsh S, Li B, Kim E, Sharipour A, Smith C, Chen YH, Berger R, Harper B, Zhang T, Park M, Shu M, Wu J, Xu J, Dewnani S, Sherer EC, Hruza A, Reichert P, Geissler W, Sonatore L, Ellsworth K, Balkovec J, Greenlee W, Wood HB. Rapid development of two factor IXa inhibitors from hit to lead. *Bioorg Med Chem Lett.* 2015 **25** 2321-2325.
44. Zögg T, Brandstetter H. Structural basis of the cofactor- and substrate-assisted activation of human coagulation factor IXa. *Structure*. 2009 **17** 1669-1678.



Appendix

Samenvatting

Curriculum vitae

Publication list

Dankwoord

SAMENVATTING

Bloedstelping (*hemostase*) is een complex proces dat gedreven wordt door een netwerk van moleculaire interacties waarin eiwitten een essentiële rol spelen. Hierin vormen de stollingsfactoren een cascade van sequentiële stappen waarin proenzymen (zymogenen) in actieve enzymen, meestal serine proteasen, worden omgezet. Dit proces wordt gereguleerd door in het bloedplasma voorkomende cofactoren, en door componenten die gelokaliseerd zijn aan het oppervlak van cellen in het bloed en in de vaatwand. Gedurende de afgelopen drie decennia is het netwerk van de hierbij betrokken eiwitten grotendeels in kaart gebracht. Dankzij kristallografisch onderzoek beschikken we veelal over de 3-dimensionale structuur daarvan. Dit vormt een belangrijke aanzet om de werking van deze eiwitten op moleculair niveau te begrijpen. Het netwerk van interacties binnen de hemostase is echter te complex om afdoende in kaart te brengen met behulp van starre structuren van de afzonderlijke componenten alleen. Hiervoor is ook inzicht nodig in de dynamische veranderingen die optreden als respons op de fysiologische omgeving. Hierbij gaat het niet alleen om interacties tussen de betrokken eiwitten onderling, maar ook om intramoleculaire interacties die bijvoorbeeld het verschil maken tussen een actieve en een niet-actieve conformatie. Hoe dit precies verloopt is voor de meeste stollingseiwitten nog een open vraag. In dit proefschrift exploreren we in hoeverre daarop een antwoord te vinden is door middel van *footprinting*-technieken in combinatie met biomoleculaire massaspectrometrie.

Deze centrale vraagstelling wordt verder uitgewerkt in **Hoofdstuk 1**. De in dit proefschrift beschreven eiwitten worden hierin geïntroduceerd, en de openstaande vragen nader gespecificeerd. Ook biedt dit hoofdstuk een beschrijving van technieken die gebruikt kunnen worden voor *footprinting* van intra- en intermoleculaire interacties. Deze methode maakt gebruik van chemische modificatie van specifieke aminozuren. De mate waarin deze modificatie plaatsvindt weerspiegelt de reactiviteit ervan, en kan beschouwd worden als een maat voor de toegankelijkheid van de desbetreffende aminozuren, zowel binnen één eiwit afzonderlijk, als in een complex ervan met één of meer andere eiwitten. Het patroon van de differentieel gemodificeerde aminozuren laat een zekere afdruk achter (*footprint*) die informatie geeft over de locatie van eiwit-eiwit contacten of over veranderingen in de 3-dimensionale structuur. In dit proefschrift wordt vooral gebruik gemaakt van twee verschillende *footprinting* technieken, die beide kort besproken worden. De eerste omvat labeling van primaire amines, dus van lysine zijketens, door middel van zogenaamde *Tandem-Mass-Tags* (TMTs). De tweede is gebaseerd op lokale massaveranderingen die binnen een eiwit optreden ten gevolge van de uitwisseling van waterstof met deuterium (*hydrogen-deuterium exchange*, HDX). Beide technieken maken gebruik van massaspectrometrie om de

mate van modificatie te kwantificeren.

Hoofdstuk 2 laat zien dat TMTs gebruikt kunnen worden voor chemische *footprinting* in de interactie tussen Receptor-Associated Protein (RAP) en het ligand bindende domein cluster II van Low-density lipoprotein Receptor-related Protein (LRP). De binding van dit eiwitcomplex berust op specifieke interacties tussen positief geladen lysine zijketens in RAP en negatief geladen bindingsplaatsen in LRP. Wij kozen dit complex voor ons onderzoek omdat de 3-dimensionale structuur ervan is opgehelderd, zodat onze experimentele strategie daarmee verifieerbaar was. Daarnaast staat RAP model voor vele andere liganden voor LRP, zoals een reeks eiwitten uit de hemostase. Onze hypothese was dat lysine zijketens in RAP in verminderde mate gemodificeerd zullen worden wanneer die afgeschermd zijn in het complex met LRP. Hiertoe werden voor RAP en RAP-LRP complex twee verschillende TMTs gebruikt, die met behulp van massaspectrometrie onderscheiden en ten opzichte van elkaar gekwantificeerd konden worden. Inderdaad werden met onze methode dezelfde lysine residuen (Lys256 en Lys270) gevonden als die eerder met behulp van kristallografie als cruciaal waren aangemerkt. Dit onderstreept de validiteit van onze TMT-benadering. Daarnaast vonden wij nog een aantal andere lysine residuen waarvan betrokkenheid in de vorming van het LRP-RAP complex nog niet eerder was gerapporteerd.

In **Hoofdstuk 3** is dezelfde strategie toegepast voor de analyse van chimere varianten van factor VIII (FVIII) waarin de twee C-domeinen één voor één uitgewisseld waren met die van factor V (FV). Van beide eiwitten wordt aangenomen dat de membraan-bindende C-domeinen twee aan twee georganiseerd zijn. Uitwisseling van het C1-domein of het C2-domein leidde tot een toename in reactiviteit van lysine residuen aan de domeinoppervlakken. Dit duidde op een verstoring van de paarsgewijze domeinoriëntatie, en een minder hechte associatie van het C2-domein met de andere domeinen van FVIII. Verrassenderwijze had dit weinig consequenties voor de functie van FVIII. FVIII met C2-domein van FV bleek functioneel vrijwel identiek te zijn aan normaal FVIII. Dat gold niet voor FVIII met het C1-domein van FV. De verklaring daarvoor is dat het C1-domein betrokken is bij interacties die specifiek zijn voor FVIII, zoals die met von Willebrand factor (VWF) en geactiveerd factor IX (FIXa). Expressie van deze chimeren in endotheelcellen suggereerde dat het intracellulair transport van de chimere eiwitten verstoord was: terwijl normaal FVIII samen met VWF teruggevonden werd in de karakteristieke opslagorganellen, de *Weibel-Palade bodies*, was dit vrijwel niet het geval voor de FVIII/FV chimeren. Blijkbaar heeft een meer flexibele oriëntatie van het C2-domein geen effect op extracellulaire functies van FVIII (interactie met FIXa en VWF) maar wel op intracellulair transport.

In **Hoofdstuk 4** hebben we hyaluronan-binding protein 2 (HABP2) bestudeerd. Dit is een serine protease waarvan de plaats binnen de hemostase controversieel is.

Ons onderzoek was in het bijzonder gericht op het Marburg-1 polymorfisme (G221E), omdat aangenomen wordt dat dit met een verhoogd tromboserisico geassocieerd is. Bovendien ligt dit polymorfisme binnen een voor serine protease karakteristiek segment, de zogenaamde *220-loop*, waardoor een functionele rol plausibel leek. Door middel van mutagenese en functionele karakterisering bleek dit voor HABP2 inderdaad het geval te zijn. De geactiveerde G221E variant vertoonde een ernstig verminderde activiteit, ook ten opzichte van eenvoudige peptidesubstraten. Normaal gaat activering van een serine protease gepaard met afsplitsing van het activeringspeptide van de N-terminus, waarna de nieuwe N-terminus door insertie in het protease domein met het actief centrum associeert. Voor HABP2 hebben wij dit onderzocht door met een biotine-label de N-terminus van het katalytisch domein te modificeren, en dit vervolgens te kwantificeren door middel van massaspectrometrie. Hierbij bleek de N-terminus van de geactiveerde G221E variant veel reactiever te zijn dan die van normaal HABP2. Dit duidt op een verhoogde toegankelijkheid van de N-terminus voor biotinylering en dus op een verminderde insertie daarvan in het katalytisch centrum. Het feit dat insertie van de N-terminus bemoeilijkt wordt door de G221E mutatie suggereert een directe relatie tussen de *220-loop* enerzijds, en het actief centrum en de N-terminus anderzijds. *Footprinting* van de N-terminus lijkt dus een relatief eenvoudig uitleessysteem te zijn voor de transitie van zymogeen in actief enzym. Dit thema hebben wij verder uitgewerkt in de volgende hoofdstukken.

In **Hoofdstuk 5** is de activering van factor IX (FIX), factor X (FX) en protrombine gevolgd aan de hand van onze chemische *footprint* benadering. Op verschillende tijdstippen tijdens de omzetting van proenzym in enzym werd met behulp van TMTs en massaspectrometrie de reactiviteit van specifieke lysine residuen kwantitatief vergeleken. Tijdens de activering van protrombine namen sommige lysine residuen toe in reactiviteit, terwijl andere juist minder reactief werden. Vooral lysines rond de *220-loop* verloren aan reactiviteit. Ditzelfde trad op tijdens de activering van FX. Echter, bij activering van FIX bleven deze veranderingen uit; opmerkelijk genoeg vertoonde de lysine residuen in de *220-loop* zelfs enige toename in reactiviteit. De ruimtelijke veranderingen rond de *220-loop* die optreden tijdens de omzetting van zymogeen in actief enzym zouden, net als bij HABP2, samen kunnen hangen met de insertie van de N-terminus in het protease domein. Dit onderzochten wij voor de volledig geactiveerde stollingsfactoren met behulp van dezelfde TMT methode. De reactiviteit van primaire amines (lysine zijketens en de N-terminus) werd vergeleken in aan- en afwezigheid van een irreversibele remmer in het actief centrum. Voor trombine vonden wij dat reactiviteit van de N-terminus ongeveer 10-maal minder was wanneer de substraatbindingsplaats bezet was. Voor geactiveerd factor IX en FX (FIXa en FXa) was dit verschil iets minder geprononceerd. Ook traden veranderingen op in de *220-loop*. Voor trombine en FXa waren deze relatief klein. In FIXa daarentegen, nam de

reactiviteit van lysines in de *220-loop* sterk af. Deze resultaten duiden op een subtiel samenspel tussen de *220-loop*, de substraatbindingsplaats, en de N-terminus van het protease domein. Verder lijkt dit mechanisme in FIXa gecompliceerder te zijn dan in FXa of trombine.

In **Hoofdstuk 6** is de rol van de *220-loop* van FIXa verder onderzocht. Daarbij lag het accent op de positief geladen residuen hierin (Lys222 en Lys224), die in het voorgaande hoofdstuk veranderingen in reactiviteit lieten zien, en op de negatief geladen residuen Glu217 en Glu219. Mutagenese resulteerde in varianten met alanine in positie 217, 219 of 224, en in de dubbelmutant 222+224. Functionele analyse toonde aan dat elk van deze mutanten enzymatisch defect was. Dit gold zowel voor de reactiviteit ten opzichte van een klein tripeptidesubstraat, als voor het fysiologische substraat FX. Alle mutanten werden gestimuleerd door FVIII, de natuurlijke cofactor van FIXa, maar bleven ver beneden de activiteit van normaal FIXa. Dit gold vooral voor de substitutie in positie 217, 224 en 222+224, en wijst op een intrinsiek defect in het katalytisch mechanisme zelf, en niet op een FVIII-afhankelijke interactie. Hetzelfde, zij het in mindere mate, gold voor de mutant E219A. *Footprinting* van de N-terminus liet zien dat de varianten met substituties in positie 217, 224 of 222+224 reactiever waren dan normaal FIXa, net als de HABP2 variant G221E in Hoofdstuk 4. De E219A variant daarentegen, vertoonde verminderde reactiviteit van de N-terminus, en een hogere reactiviteit van Lys222/Lys224 in de *220-loop*. Dit wijst erop dat aminozuur E219, hoewel belangrijk voor FIXa functie, toch een andere rol speelt dan de overige geladen residuen in de *220-loop*. Om de activerings-afhankelijke veranderingen in het serine protease domein van FIX verder te karakteriseren hebben we gebruik gemaakt van HDX. Een voordeel hiervan is dit informatie oplevert die, anders dan de TMT-methode, niet beperkt blijft tot lysine-bevattende peptiden, en ook volledig complementair is aan onze mutagenese strategie. Vergelijking van het protease domein van FIXa met FIX liet geen significante verschillen zien. Blijkbaar heeft FIXa, ondanks de proteolytische omzetting in enzym, toch vooral een zymogeen-conformatie. Dit veranderde volledig bij bezetting van de substraatbindingsplaats met een irreversibele remmer. Prominente veranderingen traden op in niet alleen de *220-loop* en bij de N-terminus, maar ook in enkele andere structuurelementen aan het oppervlak. Deze gegevens ondersteunen het concept dat de activiteit van FIXa gereguleerd wordt door een complexe, allostere relatie tussen de *220-loop* en insertie van de N-terminus. Daarnaast lijkt FIXa, meer dan in FXa of trombine, ook afhankelijk van bezetting van de substraatbindingsplaats voor het bereiken van een actieve conformatie.

In **Hoofdstuk 7** worden de bevindingen van voorgaande hoofdstukken samengevat en bediscussieerd. De rol van de *220-loop* in FIXa ten opzichte van FXa en trombine heeft daarin een prominente plaats. Ook de specifieke verdiensten en beperkingen van de *footprinting* benadering met TMTs en HDX worden geëvalueerd.

CURRICULUM VITAE

Eduard Ebberink werd op 7 juni 1985 geboren te Eindhoven. In 2003 behaalde hij zijn VWO diploma aan het Eckartcollege te Eindhoven. Vervolgens studeerde hij aan de Technische Universiteit Eindhoven waarin hij in 2010 een bachelor diploma behaalde in Biomedische Technologie en in 2012 een master diploma in *Biomedical Engineering*. Tijdens zijn bachelor doorliep hij een interne stage binnen de Organische Chemie onder leiding van dr. P.Y.W. Dankers en dr. W.P.J. Appel. Tijdens zijn stage zijn verschillende soorten supramoleculaire polymeren vermengd met peptides en onderzocht als scaffold voor niercellen. Gedurende de master heeft hij een afstudeerproject afgerond onder leiding van prof.dr. M. Merkx en dr. L.H. Lindenburg. Daar werkte hij aan de interactie tussen twee fluorescente eiwitten voor gebruik in genetisch gecodeerde *Förster resonance energy transfe*r (FRET) sensoren. Het afstudeerproject werd opgevolgd door een externe stage aan Imperial College Londen te Engeland in de groep van prof.dr. G.A. Rutter en dr. G. Meur. Daar werd onderzoek gedaan naar apoptose inductie door glucose en staurosporine in pancreatische beta cellen met behulp van genetisch gecodeerde FRET sensoren voor caspase 3 en zink. Het onderzoek naar de FRET sensoren werd vervolgens gecontinueerd als onderzoeksassistent in de groep van prof.dr. M. Merkx voor driekwart jaar. Vanaf november 2012 werd het promotieonderzoek aan Sanquin in de afdeling Plasma Eiwitten gestart, waaruit dit proefschrift is voortgekomen. Het onderzoek werd begeleid door prof.dr. K. Mertens en prof.dr. A.B. Meijer. Vanaf mei 2017 is hij werkzaam als post-doctoraal onderzoeker aan École Polytechnique Fédérale de Lausanne te Zwitserland, in een project geinitieerd door prof. dr. B. Fierz en prof.dr. P. Gönczy naar post-translationele modificaties van tubulines.

PUBLICATION LIST

Lee E, Dykas DJ, Leavitt AD, Camire RM, Ebberink EH, García de Frutos P, Gnanasambandan K, Gu SX, Huntington JA, Lentz SR, Mertens K, Parish CR, Rezaie AR, Sayeski PP, Cromwell C, Bar N, Halene S, Neparidze N, Parker TL, Burns AJ, Dumont A, Yao X, Chaar CIO, Connors JM, Bale AE and Lee AI. **Whole-exome sequencing in evaluation of patients with venous thromboembolism.** Blood Advances 2017 1:1224-1237

Stavenuiter F, Ebberink EH, Mertens K, Meijer AB. **Role of glycine 221 in catalytic activity of hyaluronan-binding protein 2.** J Biol Chem. 2017 292(15):6381-6388.

Ebberink EH, Bouwens EA*, Bloem E, Boon-Spijker M, van den Biggelaar M, Voorberg J, Meijer AB, Mertens K.* **Factor VIII/V C-domain swaps reveal discrete C-domain roles in factor VIII function and intracellular trafficking.** Haematologica. 2017 102(4):686-694.

Bloem E, Ebberink EH*, van den Biggelaar M, van der Zwaan C, Mertens K, Meijer AB.* **A novel chemical footprinting approach identifies critical lysine residues involved in the binding of receptor-associated protein to cluster II of LDL receptor-related protein.** Biochem J. 2015 468(1):65-72.

Lindenburg LH, Hessels AM, Ebberink EH, Arts R, Merkx M. **Robust red FRET sensors using self-associating fluorescent domains.** ACS Chem Biol. 2013 8(10):2133-9.

**authors contributed equally*

DANKWOORD

Onderzoek kan vrij abstract zijn maar het bestaat uit mensenwerk met veel samenwerking. Daarom wil ik hier graag de ruimte nemen om iedereen te bedanken voor hun hulp en tijd tijdens mijn werk bij Sanquin en in Amsterdam.

Beste Koen, graag wil ik je bedanken voor al je hulp en begeleiding gedurende dit promotieonderzoek. Ik heb enorm veel geleerd en genoten van alle meetings, brainstorms, werkbesprekingen, e.d.. Het blijft natuurlijk wetenschap dus het was af en toe moeilijk, maar het was zeker ook een genoegen om dit project te mogen doen. Het was een eer om door jou begeleid te worden en daarbij getuige te kunnen zijn van jouw haarscherp inzicht, biologisch redeneren, analyseren en interpreteren van resultaten en het goed opschrijven van onderzoek.

Beste Sander, ook jou wil ik graag bedanken voor alle hulp en begeleiding. Voordat ik bij Sanquin kwam wist ik vrijwel niks van de mass spec. Ik heb ontzettend veel van je geleerd en ben je dankbaar voor de mogelijkheden die je me geboden hebt, zoals het beoefenen van HDX. Ik bewonder je enthousiasme, het razendsnel interpreteren van resultaten en bedenken van experimenten en ideeën. Ik wil je bedanken voor je enorme inzet en hulp. Het was een enorm plezier om te mogen werken binnen de proteomics groep.

E poi, vorrei ringraziare le mie paraninfe. Fiamma! So much has happened in those 4.5 years. I can't thank you enough for all the help and support given those years. From idiotic couscous salads to important presentations you were willing to help me. By pure coincidence I ended up in Orteliuskade and we became neighbours. A starting point for a great time in Amsterdam, I will always cherish those great times, the dinners, parties, the dancings and even the (involuntary) supporting of team Italy. I am so grateful we became such close ~~colleagues~~ friends. In good times and in bad you were there, thank you so much! Dear Nadia, you came a bit later during my PhD to Sanquin and in a short time you became a close friend and big help. I want to thank you for all the fun discussions, scientific and non-scientific, and help towards the end of the project. The late Sunday afternoon FXa generations assays became fun to do, irrespective of our difficult-to-match music styles. I thank you for all the support also during those (rare?) moments of me complaining. I couldn't have managed without your help and I couldn't have wished for a better lab-bench partner and friend. Dear *paraninfe*, both your energetic passions for science are an example for me!

Al snel tijdens mijn promotie verhuisde we naar het Y-gebouw. Daar kwam ik terecht in Y420 lichtelijk verstopt in de hoek maar met schitterend uitzicht op Sanquin en het NKI. Maartje, ik wil je bedanken voor alle tips, adviezen en input gedurende mijn project, tijdens de werkbesprekingen of als ik met verwarrende resultaten de kamer inliep. En na mijn verhuizing vanuit het warme zuiden naar het kille noorden (valt mee) was er in ieder geval iemand die de schoonheid van een worstenbroodje en PSV CL overwintering kon waarderen! Herm-Jan, dank voor al je hulp, menigmaal heb ik gebruik gemaakt van je uitgebreide naslagwerk over de coagulatiecascade, en ik heb genoten van de gesprekken over de gang van zaken op Sanquin en de rest van de wereld. Eva, ook wij hebben heel wat afgekeuveld en het was goed om daarbij vanuit een frisse dokters blik context te krijgen, dank! Arjan, je kwam wat later als kantoorgenoot naar Y420 met een goede dosis humor en je hebt me getriggerd om weer eens een poging te wagen tot programmeren, (Matlab staat inmiddels weer op de laptop oei!). Beste Y420 kamergenoten, ik wou jullie bedanken voor alle gebrachte koffies en de zeer prettige werkomgeving.

I have made a lot of great friends during my years in Amsterdam. Dear Nicoletta, I am so glad to have met you, also together with Fiamma as neighbour at Ortelius. We had such tremendous fun in New York! Yes yes 'the metropolitan museum glove incident', but come on, zero dollar admittance! The dancings, the dinners I enjoyed it enormously and it made my time in Amsterdam. Also when things hit the fan, you were always there to talk and help. Up till today, although you are a bit further away now in the USA. However, friendship doesn't count miles (got that from the internet). Ook wanneer je naar Kopenhagen verhuist. Esther jouw werk met TMTs was het startpunt voor mijn onderzoek. Je hebt me enorm geholpen tijdens dit project, je uitermate meticuleuze manier van werken, daar kan ik zeker nog aan tippen. Van avondjes uit tot gezamenlijk RAP zuiveren het was een mooie tijd! Dank voor alle hulp, adviezen en onze vriendschap. Sten, team Ortelius, I want to thank you too for all advices and support, the crazy carnaval and musical/lyrical analyses. Aleksandra, you were just finishing and getting ready to move to NY. I would like to thank you for the great time we had in Nowy Jork and at Sanquin. Dorothee, Claudia, Bieuwke, Kathinka, Magdalena, Javier I will like to thank you all for the valuable input especially in the beginning of my Amsterdam time. Javi, I appreciated your warnings when my dutch 'politeness' would collide with the southern politeness. I should have listened!

Robin, we hebben heel wat geredekaveld (dat leek me nou een mooi woord) waarbij vooral veel gelachen, van Kanariestraat, tot babysitten (ja ja babysitten(!) op de kleine Jasper). Van Elburg tot het Orgelmuseum Utrecht, en het Zoutmuseum in Liverpool (USA) we hebben veel meegemaakt. Hoe gaat het met de mondharp? Het was een

genoegen om met je te kunnen praten over wetenschap en daarbuiten. Hopelijk kunnen we goede vertellingen blijven toevoegen. Fabian, we begonnen op de dezelfde dag aan ons promotieonderzoek. Nog in de 'kelder' in het oude gebouw en vanaf die dag hebben we vaak even kunnen ventileren over ons promotieonderzoek. Jouw invulling van Bon Jovi blijft ongenaakbaar! Met zijn drieën begonnen we ongeveer tegelijkertijd en ook eindigen we nu in hetzelfde jaar, een soort van 'class of 2017'.

The new generation, Malgosia, I would like to thank for your help and fruitful discussions inside and outside of the lab. Hopefully some of the hard HDX work will pay off in the end! You brought a much appreciated, fresh joy into the lab and don't underestimate the importance of you safeguarding sublime fm and the ability to relativate and put things in perspective on difficult days. Bart, Annemarie, Maaike S, Ellie, Eelke, Johana, Maaike R, Esmee, you all came a bit later to the lab and we had great times and talks, from the lunches, to the garden parties, birthday celebrations, to conferences like the NVTH and ISTH. There was always a great atmosphere and willingness to help each other, for which I would like to thank you all very much.

Over bereidheid tot helpen gesproken, ik wil graag het team van technici bedanken. Zij hebben altijd geholpen wanneer ik verloren was in mijn protocol of er een verassing kwam opduiken tijdens een experiment. Mariëtte, jij was de eerste die mij hielp in het lab. Je hebt mij alles geleerd over de assays en zuiveringen, van de meng/roltechniek voor de platewasher buffer tot het maken van FIX mutanten. Ik kan me nog herinneren dat ik, na onze eerste zuivering, de kostbare FIX pool heb laten vallen na dialyse (oei!), volgens mij ben ik toen 60% verloren. Dankzij jou heb ik nieuw FIX kunnen zuiveren en ligt er uiteindelijk nu nog puur FIX in de vriezer. Ik bewonder je geduld en altijd parate bereidheid tot hulp. Carmen, net als Mariëtte, een rots in de lab-branding, ik dank je voor het altijd klaar staan zelfs over whatsapp als de MS weer rood begon aan te slaan. Floris, dank voor alle hulp vooral met de HDX en MS. Ik had geen idee wat een bahco of ringsleutel was, gelukkig wist je meteen raad wat te doen zo gauw de HDX, MS of PEAKS weer haperde. Paul en Marleen, af en toe ging het weer mis, lekkende kolom of kapotte pomp dank voor alle hulp in lab! Menno, dank voor alle interessante discussies zowel wetenschappelijke als daarbuiten.

Beste Jan, dank voor alle inbreng en getoonde interesse gedurende mijn project. Tijdens werkbesprekingen maar ook daarbuiten, bijvoorbeeld in de rij voor het Y3 koffie apparaat, wat tevens ook mooie onderonsjes opleverde! Het was een eer om jouw inaugurele rede bij te mogen wonen. Dank!

Bij de kennismaking kwam als een van de eerste slides het Philips-stadion opdoemen, en dat in Amsterdam! Beste Joost, ik wou je bedanken voor je adviezen, wetenschappelijke input en de korte onderonsjes. En voor je passie voor het verspreiden van onderzoek, dankzij jou zijn er online zowaar bewegende beelden beschikbaar dat ik daadwerkelijk in het MS hok aan het werk ben. Ik heb mijn best gedaan om zo zorgelijk mogelijk naar de HDX te kijken. Kamran, Tom, Ingrid, Arnoud, dank voor alle hulp in en rond het lab!

Ruben, dank voor alle goede adviezen en mooie gesprekken, avondjes uit en natuurlijk het zijn van mijn #1 fan! Een high potential post-doc om naar op te kijken. Anastasia and Maryam thanks for the great times! Ivan, maintenain je dois apprendre francois et ce ne pas facile! I would like to thank you for all your valuable input and nice afterwork discussions and cyclings after a few pilsies. And in the wake, Laura thanks for the great times and fun, I will never say 'Catalunians' again! Josse, na mij de eerste die de geduchte taak op zich nam om te gaan HDXen. Dank voor alle hulp, frisse invalshoeken en ideeën voor mijn onderzoek. Henriët, Marjon, Sjoert, dank voor alle hulp en gezelligheid!

Beste (voormalige) MCBers en co, Lilian, Simon, Mark, Dion, Jeffrey, Carlijn, Anne, Rivellino dank voor de nodige 'breaks' van voornamelijk niet-wetenschappelijke aard. Doen de vissen het nog? Mar, I would like to thank you for the nice conversations, valuable advices and of course the great laughs we had! Stefan, Yvonne, Kalim, Michel, dank voor alle mooie gesprekken tijdens borrels of tijdens het voorbij komen op de gang. Beyond the Y-building I had the pleasure to meet a lot of other people and make new friends. Tati (yes yes friends) I would like to thank you for your crazy fun ways and advices for my abroad move. Peter-Paul, Felix, Anna, Aurelie, Tamara, Francesca, thanks for the BBQ's, 90's parties and fun evenings. Everybody else from the U-building (sorry I still can't distinguish between all of these departments) thanks for inviting me from across the institute and the good times we had at Sanquin/ Amsterdam.

Beste Marion, Gerty, Wanda dank voor alle administratieve hulp (en nog mijn excuses voor alle late WSBO formulieren).

Corrie, Johanna en Suzanne dank voor alle hulp met bestellingen!

Mijn lieve vrienden uit Eindhoven en omstreken, de biki en aanverwanten, toppers, dank voor de good times tijdens mijn promotie. Van gestrande fietstochten in Eersel,

avondjes in het Philips-stadion tot excursies bij het Gardameer en IJmuiden, dank voor de dolle boel. Tjonge, wat een gouden tijd in Eindhoven en hopelijk komen er nog vele verhalen bij, vooral als ik weer wat meer tijd heb. Rene, dank voor de 'artist in residence' weekendjes te Tongelre!

Tijdens dalmomenten, hamerden mijn ouders er steeds op om maar te kijken hoe de hazen lopen en vooral om mijn zegeningen te tellen. Twee zegeningen zijn zonneklaar, Mams, Paps dank voor alle steun en hulp. Zonder af en toe een belletje en de mogelijkheid om te kunnen relativeren was dit nooit gelukt. Naast het onderzoek gebeurt er natuurlijk van alles en jullie stonden daarbij altijd voor me klaar. Laurens, denk ik soms een goeie grap te maken word ik glazig aangestaard, gelukkig kon ik met jou altijd de infame Ebberink-Roosen-humor delen. Het is jammer dat de tijden op de achterbank met grappen over Fred Duchateux en Meppel (dreun!) voorbij zijn, maar over de wifi komen ze toch nog goed door. Naast humor denken we over veel hetzelfde en ben ik blij en trots zo'n grote broer te hebben. Jolien en de kleine koters, Tess en Roan, een bezoekje of etentje was altijd verfrissend en goed om me uit de onderzoeks-slop te geraken. Na mijn Amsterdam/Lausanne strapatsen (een typisch Ebberink woord lijkt mij) zou ik wat vaker op bezoek moeten komen! Het leven van een maffe oom gaat niet altijd over rozen, maar het thuisfront maakt het een heel stuk lichter, dank!