Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Mutation of surface residues to promote crystallization of activated factor XI as a complex with benzamidine: an essential step for the iterative structure-based design of factor XI inhibitors

Lei Jin,* Pramod Pandey, Robert E. Babine, David T. Weaver, Sherin S. Abdel-Meguid and James E. Strickler

Daiichi Asubio Medical Research Laboratories LLC, USA

Correspondence e-mail: leijin05@yahoo.com

Received 10 June 2005
Accepted 28 July 2005

PDB References:
rhFXI$_{370\text{-}607}$-S434A,T475A, K437A–benzamidine, 1zhm, r1zhmsf; rhFXI$_{370\text{-}607}$-S434A, T475A,K505A–benzamidine, 1zhp, r1zhpsf; rhFXI$_{370\text{-}607}$-S434A,T475A,C482S, K437A–benzamidine, 1zhr, r1zhrsf.

Activated factor XI (FXIa) is a key enzyme in the amplification phase of the blood-coagulation cascade. Thus, a selective FXIa inhibitor may have lesser bleeding liabilities and provide a safe alternative for antithrombosis therapy to available drugs on the market. In a previous report, the crystal structures of the catalytic domain of FXIa (rhFXI$_{370–607}$) in complex with various ecotin mutants have been described [Jin *et al.* (2005), *J. Biol. Chem.* **280**, 4704–4712]. However, ecotin forms a matrix-like interaction with rhFXI$_{370–607}$ and is impossible to displace with small-molecule inhibitors; ecotin crystals are therefore not suitable for iterative structure-based ligand design. In addition, rhFXI$_{370–607}$ did not crystallize in the presence of small-molecule ligands. In order to obtain the crystal structure of rhFXI$_{370–607}$ with a weak small-molecule ligand, namely benzamidine, several rounds of surface-residue mutation were implemented to promote crystal formation of rhFXI$_{370–607}$. A quadruple mutant of rhFXI$_{370–607}$ (rhFXI$_{370–607}$-S434A,T475A,C482S,K437A) readily crystallized in the presence of benzamidine. The benzamidine in the preformed crystals was easily exchanged with other FXIa small-molecule inhibitors. These crystals have facilitated the structure-based design of small-molecule FXIa inhibitors.

1. Indroduction

Hemostasis is a complex process involving a delicate balance between the activation of a cascade of serine proteases to bring about thrombosis and the release of soluble inhibitors to limit their action. Disruption of the balance between activation and inhibition can lead to life-threatening conditions such as deep-vein thrombosis or arterial thrombosis, leading to stroke, acute pulmonary embolism and myocardial infarction. Heparin and coumadin (warfarin) are widely used to prevent thrombosis; they interfere at multiple points in the coagulation cascade and have severe bleeding liabilities. Therefore, their dosage has to be adjusted for each patient and the extent of their inhibition of clot formation must be monitored on a regular basis during therapy. Heparin can also induce thrombocytopenia, an autoimmune disorder, in a small number of patients. Heparin-induced thrombocytopenia is fatal unless immediately and aggressively treated. In attempts to develop better antithrombotic agents, many inhibitors have been developed for factor Xa and for thrombin. Both factor Xa and thrombin are key enzymes at the final steps of the coagulation cascade; the potential bleeding liability associated with inhibiting these enzymes has not yet been fully established (Hirsh *et al.*, 2005).

FXI plays an essential role as an amplifier of the coagulation response (Walsh, 2001). Activated FXI (FXIa) has been found to be elevated in patients surviving acute myocardial infarction, suggesting an involvement of this factor in acute coronary disease. Recent reports implicate a role for FXIa in maintaining clot integrity (Bouma & Meijers, 2000; Minnema *et al.*, 1998), while individuals with FXI deficiency do not show severe bleeding problems (Ragni *et al.*, 1985). Additionally, an anti-human FXI antibody did not prevent thrombus initiation, but significantly reduced intraluminal thrombus growth in a baboon thrombosis model (Gruber & Hanson, 2003). The anti-FXI antibody prolonged the partial thromboplastin time (a measure directly involving FXIa), but did not affect prothrombin time (a measure involving FVIIa and tissue factor) and most importantly did not prolong bleeding time. This poses FXIa as an attractive target for developing novel anticoagulant agents that may be advantageous over warfarin, heparin, low-molecular-weight heparins and thrombin inhibitors for prophylaxis and treatment of thromboembolic disorders.

FXI is a 160 kDa disulfide-linked homodimer consisting of two identical subunits each containing 607 amino acids and is activated by cleavage after residue 369 of each of the two subunits, producing a protein that contains two heavy chains and two light chains. The heavy chain contains four tandem repeat sequences, designated 'apple' domains, containing binding sites for platelets (Baglia *et al.*, 1995), heparin (Ho *et al.*, 1998) and many proteins, such as thrombin (Baglia & Walsh, 1996), high-molecular-weight kininogen (Baglia *et al.*, 1992), factor XIIa (Baglia *et al.*, 1993) and glycoprotein Ibα (Baglia *et al.*, 2004). The light chain, consisting of residues 370–607, contains a catalytic domain belonging to the trypsin-like serine protease family, highly homologous to the catalytic domains of other factors in the blood-coagulation cascade, such as FIXa, FXa, FVIIa and thrombin. There is a single disulfide bond connecting the heavy chain and the light chain in each subunit.

We have reported the crystal structures of the FXIa catalytic domain (rhFXI$_{370–607}$) in complex with various unnatural ecotins (Jin *et al.*, 2005), a protease-inhibitor protein found in *Escherichia coli* containing 142 amino acids (Chung *et al.*, 1983). Although the structures provided three-dimensional information for the catalytic domain of FXIa, as well as substrate-like interactions, these crystals were not suitable for the iterative process of structure-based ligand design. Here, we present the crystal structure of rhFXI$_{370–607}$ with benzamidine, a small-molecule serine protease inhibitor. As benzamidine is a weak inhibitor, it can be easily exchanged with other small-molecule ligands. Significant effort was made to obtain crystallizable rhFXI$_{370–607}$ *via* mutations of surface residues. A quadruple mutant of rhFXI$_{370–607}$ was identified that produced stable high-resolution crystals with benzamidine and could be readily exchanged with other small-molecule inhibitors.

We use two numbering systems for rhFXI$_{370–607}$ in this paper. For the mutagenesis studies, we use FXIa sequence numbers. For the structure discussion, we converted the FXIa sequence number to chymotrypsin sequence numbers (in italics) for easy structural comparison with other members of the trypsin-like serine protease family. The corresponding residue numbers for the FXI sequence and chymotrypsin numbering systems are listed in the supplementary material of Jin *et al.* (2005).

2. Experimental procedures

2.1. Recombinant methods and mutagenesis

The cDNA of FXI was kindly provided by Professor Peter Walsh of Temple University. The DNA encoding the catalytic domain (residues 370–607) was amplified by polymerase chain reaction (PCR). *Xho*I and *Not*I restriction sites were included in the PCR primers. The digested PCR product was cloned into *Xho*I–*Not*I sites of the *Pichia* vector pPICZa A (Invitrogen). This plasmid was used to generate various mutants by oligonucleotide-directed site-specific mutagenesis using QuikChange reagents (Stratagene). Plasmid DNAs were linearized with the restriction enzyme *Sac*I prior to transformation. *Pichia pastoris* strain X-33 was used for expression and routinely grown in buffered glycerol complex medium; buffered methanol complex medium was used for induction of protein expression.

2.2. Protein expression and purification

Recombinant human FXI$_{370–607}$ and mutants thereof were purified by modifications of a previously published method (Jin *et al.*, 2005). *Pichia* conditioned medium was concentrated sevenfold to tenfold using a Pell concentrator (10 000 NMWCO membrane) and subjected to two rounds of fivefold dilution and concentration with 20 m*M* Tris–HCl pH 7.4. The concentrated diafiltered medium was either used immediately or frozen at 193 K until use. The medium was thawed at room temperature overnight and then applied onto a 1.6 × 20 cm SP-Sepharose column (Amersham Biosciences) equilibrated with 20 m*M* Tris–HCl pH 7.4 at a flow rate of 8–10 ml min^{-1}. After loading, the column was washed with the same buffer until the A_{280} of the effluent returned to baseline. The column was then washed with 0.075 *M* NaCl in the same buffer and rhFXI$_{370–607}$ was eluted with 0.2 *M* NaCl, 20 m*M* Tris–HCl pH 7.4. (The K505A mutant was retained less strongly on this column and eluted in the 0.075 *M* NaCl wash.) The eluted protein was diluted with an equal volume of 20 m*M* Tris–HCl, pH 7.4 and applied onto a 5 ml Heparin HiTrap column (Amersham Biotech) equilibrated in 0.1 *M* NaCl, 20 m*M* Tris–HCl pH 7.4. Bound rhFXI$_{370–607}$ was eluted with a linear 50 ml gradient from 0.1 to 0.5 *M* NaCl in the same buffer. The eluted enzyme was collected and applied onto a 2.5 × 80 cm Superdex 75 column equilibrated in 0.2 *M* NaCl, 20 m*M* Tris–HCl pH 7.4 at a flow rate of 4 ml min^{-1}. The enzyme eluted at a position consistent with a 30 kDa protein. Aliquots of the purified protein were frozen at 193 K for crystallization experiments.

2.3. FXIa assays

Native human FXIa (Haematological Technologies) and rhFXI$_{370–605}$ were assayed using a modified version of the standard FXIa peptide substrate S2366, *i.e.* pyroGlu-Pro-Arg-*p*-nitroanilide, in which the *p*-nitroaniline reporter was replaced with aminomethylcoumarin (AMC). The assay was carried out at 303 K in 20 m*M* Tris–HCl pH 7.4, 150 m*M* NaCl, 0.02% Tween 20 with a substrate concentration of 50 µ*M* and an enzyme concentration of 0.25 n*M*. Under these conditions, the assay was linear for at least 60 min.

2.4. Crystallization and data collection for the mutant rhFXI$_{370–607}$–benzamidine complex

Benzamidine was added to each of the rhFXI$_{370–607}$ mutants to a concentration of 1 m*M*. The complex was then concentrated to about 20 mg ml^{-1} with buffer containing 0.2 *M* NaCl, 20 m*M* Tris–HCl, 1 m*M* benzamidine pH 7.4 at 277 K. Crystallization of the rhFXI$_{370–607}$–benzamidine complex was carried out using the hanging-drop vapour-diffusion method. Cubic crystals were grown in 2 µl drops containing equal volumes of rhFXI$_{370–607}$–benzamidine complex solution and precipitant solution consisting of 2.0 *M* ammonium sulfate and 0.1 *M* Tris–HCl pH 8.5 and equilibrated against 300 µl precipitant solution at 283 K.

Diffraction data for the rhFXI$_{370–607}$-S434A,T475A,K437A–benzamidine crystal were measured at beamline X12C of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory. Data collection for the rhFXI$_{370–607}$-S434A,T475A,K505A–benzamidine crystal and the rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine crystal was performed on an R-AXIS IV^{++} imaging-plate detector with an RU-H3R generator and an X-stream 2000 system (Rigaku/MSC) at 123 K. All crystals were soaked briefly (less than 10 s) in cryoprotection solution consisting of crystallization solution with 20% glycerol before being flash-frozen in liquid nitrogen. All data were processed using *HKL*2000 (Otwinowski & Minor, 1997). The crystals belong to the cubic space group *I*23, with one rhFXI$_{370–607}$ molecule in the asymmetric unit. Summaries of the unit-cell parameters and statistics of diffraction data are listed in Table 1.

2.5. Structure solution and refinement

The initial phases for the rhFXI$_{370–607}$-S434A,T475A,K437A–benzamidine structure were obtained by molecular replacement using *AMoRe* (Navaza, 2001). The search model was one rhFXI$_{370–607}$ molecule from the tetrameric structure of the rhFXI$_{370–607}$–ecotinM84R complex (Jin *et al.*, 2005). To avoid model bias, all the surface loops, the side-chain atoms beyond C$^{\beta}$ and water molecules were removed. A clear solution was obtained using 20–5 Å data from the rhFXI$_{370–607}$-S434A,T475A,K437A–benzamidine crystal, with a correlation coefficient of 0.506 and an *R* factor of 34.8%. Structure refinement was carried out using *CNX* (Accelrys) and model building was performed in *QUANTA* (Accelrys). The coordinates of the rhFXI$_{370–607}$-S434A,T475A,K437A–benzamidine structure were used subsequently for refinement and model building against data from the rhFXI$_{370–607}$-S434A,T475A,K505A–benzamidine crystal and the rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine crystal. Detailed refinement and structure statistics are listed in Table 1.

2.6. Ligand-soaking procedure for rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine crystals

A small amount of dry compound was added to a drop of mother liquor to make a saturated solution on the same cover slide on which the benzamidine crystals were grown. The ligand-containing drop was equilibrated for more than 15 min over the reservoir solution before a benzamidine crystal was transferred over. The crystal was soaked for a couple of hours to days at 283 K before data collection on an R-AXIS IV^{++} imaging-plate detector.

3. Results and discussion

3.1. Search for the optimum mutant of rhFXI$_{370-607}$ protein

Our first rhFXI$_{370–607}$ structure was a complex with the protein protease inhibitor ecotinM84R using enzymatically deglycosylated wild-type rhFXI$_{370–607}$ (Jin *et al.*, 2005). That structure and the structures with ecotins having the P5 to P2′ residues mutated to the FXIa substrate sequence provided substrate-like binding information; however, the crystals of the complex were not suitable for the iterative structure-based ligand-design process. The ecotin dimer and the two rhFXI$_{370–607}$ molecules form a tetrameric complex in the rhFXI$_{370–607}$–ecotinM84R structure. It is not feasible to displace ecotin with a small-molecule ligand in these crystals. The catalytic domain of FXIa has two N*X*S/T glycosylation sites. The protein-proximal *N*-acetylglucosamine (GlcNAc) portion of oligosaccharide side chains cannot be removed by endoglycosidase H. In the rhFXI$_{370–607}$–ecotinM84R structure (PDB code 1xx9), there is one GlcNAc molecule visible, attached to Asn473 (*Asn113*; residue numbers in parentheses and in italics correspond to the chymotrypsin sequence-numbering system) in one of the two FXIa molecules in the tetrameric FXIa–ecotin complex. We made an rhFXI$_{370–607}$-S434A,T475A mutant to remove the heterogeneity that arises from the two glycosylation sites; however, this protein in complex with benzamidine only produced thin needle-like crystals. After exhaustive optimization of the crystallization conditions, rod-like crystals were obtained, but they diffracted anisotropically and were not suitable for data collection.

To further modify the protein for crystallization, we performed lysine-to-alanine mutations on rhFXI$_{370–607}$-S434A,T475A based on reports that mutation of surface residues promoted crystallization of RhoGDI (Derewenda, 2004; Longenecker *et al.*, 2001, 2003). The hypothesis is that surface residues with high conformational entropy, such as lysines, are likely to impair the formation of protein–protein contacts in crystal lattices. Replacing lysines with alanine created epitopes that could be incorporated into crystal contacts with lower entropic penalty, as well as removing charges to alter the protein's isoelectric point and solubility

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

	rhFXI$_{370–607}$-S434A, T475A,K437A	rhFXI$_{370–607}$- S434A, T475A,K505A	rhFXI$_{370–607}$- S434A, T475A,C482S,K437A
Diffraction data			
Resolution (Å)	50–1.96 (2.03–1.96)	50–2.7 (2.8–2.7)	30–1.73 (1.79–1.73)
Space group	*I*23	*I*23	*I*23
Unit-cell parameters (Å)	120.1	120.2	120.7
Completeness (%)	100 (100)	99.9 (100)	99.9 (99.9)
Redundancy	16.0	5.3	5.8
$\langle I\rangle/\langle\sigma(I)\rangle$	36.0 (6.7)	16.2 (3.3)	32.0 (4.1)
$R_{\rm merge}$ (%)	7.5 (33.6)	7.4 (37.5)	9.4 (41.7)
Refinement			
Data (no cutoff) (Å)	40–1.96	40–2.7	30–1.73
R factor (%)/No. of reflections	17.0/18517	20.8/7006	18.3/27136
$R_{\rm free}$ (%)/No. of reflections	19.1/2022	25.7/830	20.2/3040
R.m.s.d. in bond lengths (Å)	0.006	0.007	0.0045
R.m.s.d. in bond angles (°)	1.20	1.10	1.20
R.m.s.d. in dihedrals (°)	24.9	25.8	24.8
Mean *B* factors (Å^2)	22.5	44.7	26.5
Ramachandran plot			
Favored (%)	87.2	80.8	85.8
Additionally allowed (%)	12.8	19.2	14.2
Generously allowed (%)	0.0	0.0	0.0
Disallowed (%)	0.0	0.0	0.0

properties. We located all 16 lysine residues of rhFXI$_{370–607}$ in the structure of the rhFXI$_{370–607}$–ecotinM84R complex. Five lysines, Lys422 (*Lys63*), Lys437 (*Lys78*), Lys486 (*Lys127*), Lys505 (*Lys145*) and Lys509 (*Lys149*), were chosen for mutation to alanine individually in the rhFXI$_{370–607}$-S434A,T475A mutant. All five residues are on the surface of rhFXI$_{370–607}$, are away from the active site, are not involved in intramolecular interactions and are poorly conserved in FXIa from other species.

Two out of the five lysine mutants, K437A and K505A, in complex with benzamidine, produced new cubic crystals. The crystal of the K437A mutant diffracted to 1.96 Å resolution at beamline X12C of NSLS at the Brookhaven National Laboratory. After solving the structure of the K437A mutant, it was clear that the mutated lysine was involved in crystal packing. As shown in Fig. 1(*a*), the C$^{\beta}$ of Ala437 (*Ala78*) pointed directly to the guanidium group of Arg586 (*Arg224*) of a symmetry-related molecule (Arg586$_{\rm sym}$). Arg586$_{\rm sym}$ also formed two hydrogen bonds with Asp439 (*Asp80*) and interacted with the carbonyl O atom of Ala434 (*Ala75*) through a water molecule. From this structure, it was apparent that the long side chain and positive charge of Lys437 would interfere with Arg586$_{\rm sym}$ at this lattice-contact point. Interestingly, the crystal structure of the K505A mutant confirmed this notion.

Although the morphology of the K505A crystals was the same as that of the K437A crystals, the K505A crystals consistently diffracted to lower resolution than the K437A crystals on the same R-AXIS IV^{++} detector, after testing several K505A crystals with sizes similar to or larger than that of the K437A crystals. In the structure of the K505A mutant at 2.7 Å resolution, Lys437 is not in an extended conformation (Fig. 1*b*). When the structures of the K505A and K437A mutants were compared, the main-chain atoms around Lys437 in the K505A structure were pushed away from Arg586$_{\rm sym}$ owing to charge repulsion between Lys437 and Arg586$_{\rm sym}$ and/or steric hindrance. The distances between the hydrogen-bonding atoms also increased slightly (Fig. 1*b*). Thus, the K505A mutant does not pack as tightly as the K437A mutant at this lattice-contact point in the cubic crystals. This might be one reason for the poor diffraction quality of the K505A crystals.

Residue 505 was involved in lattice contacts in both the K437A and K505A crystal structures, close to Arg507 (*Arg147*) from a symmetry-related molecule (Arg507$_{\rm sym}$). In the K437A structure both Lys505 and Arg507$_{\rm sym}$ had disordered side chains, while in the K505A structure the smaller side chain of Ala505 allowed the ordering of Arg507$_{\rm sym}$.

Two other lysine mutants, K422A and K509A, produced only needle-like crystals that were not suitable for data collection. The K486A mutant had a very low expression level. Upon close inspection of Lys486 (*Lys127*) in the structures of K437A and K505A, Lys486 (*Lys127*) formed a salt bridge with Asp598 (*Asp236*) located on the C-terminal helix. The interaction between Lys486 and Asp598 might be important for the structural integrity of rhFXI$_{370–607}$. Re-examination of the rhFXI$_{370–607}$–ecotinM84R structure revealed that Lys486 (*Lys127*) also interacts with Asp598 (*Asp236*). Lys486 was mistakenly picked for mutation, but served as a control. Taken together, these results show that selected surface lysines, but not structurally important ones, could be safely mutated without altering the overall conformation of the enzyme. Additionally, kinetic parameters, *e.g.* the $K_{\rm m}$ and $k_{\rm cat}$ for S2366, did not differ significantly between native FXIa isolated from plasma and several of the mutants described here: rhFXI$_{370–607}$, rhFXI$_{370–607}$-S434A,T475A, rhFXI$_{370–607}$-S434A,T475A,K505A and rhFXI$_{370–607}$-S434A,T475A,C482S, K437A (data not shown.)

In the structures of the rhFXI$_{370–607}$-S434A,T475A,K437A–benzamidine complex and the rhFXI$_{370–607}$-S434A,T475A, K505A–benzamidine complex, we observed additional density that was attached to Cys482 (*Cys123*). In the full-length FXIa, Cys482 forms a disulfide bond with Cys362 in the fourth 'apple' domain. Cys362 was not present in the rhFXI$_{370–607}$ construct and as a result Cys482 was left as a free cysteine. Based on the shape of the density extending from Cys482 (*Cys123*), we suspected that glutathione (γ-glutamylcysteinylglycine, GSH), a sulfhydryl antioxidant that is ubiquitous in animals, plants and microorganisms, was covalently attached to Cys482 by a disulfide bond. GSH often attains millimolar levels inside cells, which makes it one of the most highly concentrated intracellular antioxidants. It is likely that GSH became attached to the protein during expression in *P. pastoris*. GSH fits well into the density, except that the

glutamyl portion of GSH was disordered and truncated during structure refinement (Fig. 1*c*). The GSH molecule was not only packed tightly with rhFXI$_{370–607}$ but was also involved in lattice contacts with two symmetry-related molecules as shown in Fig. 1(*c*) with electron densities labeled as Sym1 and Sym2. To further reduce the disordering at this lattice-contact point, we subsequently added the C482S mutation to remove this GSH modification site. The quadruple mutant rhFXI$_{370–607}$-S434A, T475A,C482S,K437A produces crystals of high diffraction quality reproducibly and has proven to be essential for carrying out iterative structural-based ligand design.

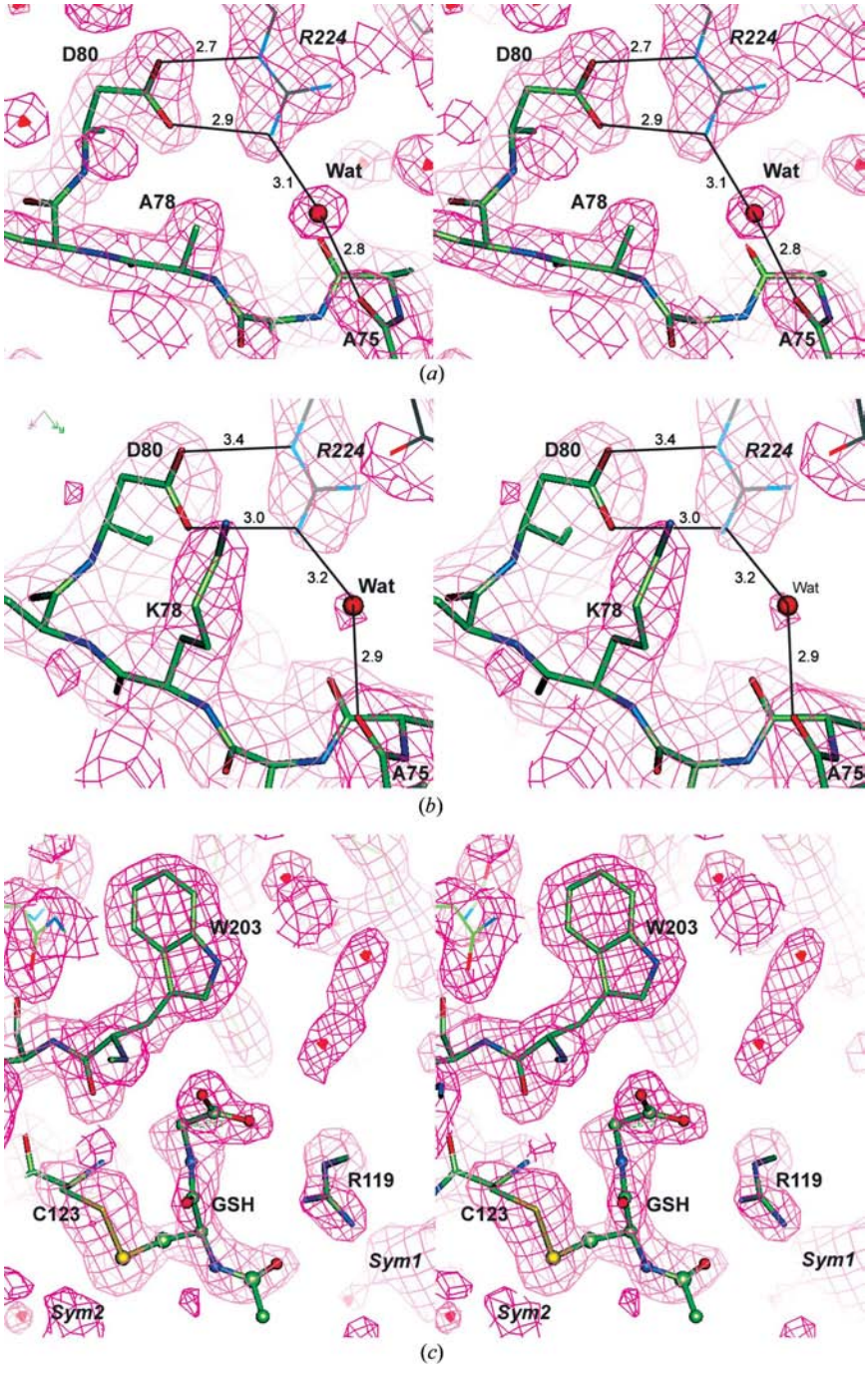

Figure 1
Lattice-contact points for the cubic crystals of rhFXI$_{370–607}$ mutants in complex with benzamidine. All maps are $2F_o - F_c$ contoured at 1σ and are presented as pink wires. Hydrogen-bonding atoms are connected by grey lines with distances labeled in Å. Water molecules are labeled as Wat. (*a*) The lattice-contact point at *Ala78* in the structure of rhFXI$_{370–607}$-S434A,T475A,K437A–benzamidine. The residues are in stick representation, with thicker green C atoms for one molecule and thinner grey C atoms for *Arg224* of a symmetry-related molecule. (*b*) The lattice-contact point at *Lys78* in the structure of rhFXI$_{370–607}$-S434A,T475A,K505A–benzamidine. The color coding is the same as (*a*). (*c*) Residues and electron-density map around *Cys123* in the structure of rhFXI$_{370–607}$-S434A,T475A,K437A–benzamidine. Glutathione (GSH) without the glutamyl portion is presented in ball-and-stick representation and forms a disulfide bond with *Cys123*. The two pieces of densities labeled Sym1 and Sym2 at the bottom of the picture belong to two symmetry-related molecules.

3.2. Structure comparison of rhFXI$_{370–607}$ in complex with benzamidine and ecotinM84R

The highest resolution data was measured from a crystal grown with the quadruple mutant of rhFXI$_{370–607}$ and benzamidine. The crystal diffracted to 1.73 Å resolution on an R-AXIS IV^{++} detector with a RU-H3R generator and blue Osmic mirrors. This structure is very similar to the 1.96 Å resolution structure of rhFXI$_{370–607}$-S434A,T475A,K437A in complex with benzamidine. The rhFXI$_{370–607}$-S434A, T475A,C482S,K437A–benzamidine structure is used for all subsequent structural comparison.

The overall structures of the rhFXI$_{370–607}$ in the benzamidine and ecotinM84R complexes are very similar. When all the C$^{\alpha}$ atoms in the two rhFXI$_{370–607}$ molecules are superimposed, the root-mean-square (r.m.s.) deviation is 0.69 Å compared with chain *A* of the rhFXI$_{370–607}$–ecotinM84R structure (0.51 Å compared with chain *B* of the rhFXI$_{370–607}$–ecotinM84R structure). Most of the structural changes are located at surface loops, residues *36–37B*, *125–133* and *143–152*, which are involved in crystal packing. When those loops are excluded from the C$^{\alpha}$ alignment, the r.m.s. deviation is 0.52 Å compared with chain *A* of the rhFXI$_{370–607}$–ecotinM84R structure (0.45 Å compared with chain *B* of the rhFXI$_{370–607}$–ecotinM84R structure). The active site of rhFXI$_{370–607}$, as in all other trypsin-like proteases, is located at the cleft between two β-barrels. When the C$^{\alpha}$ atoms of one β-barrel (residues *31–122*) are superimposed, a shift of the second β-barrel is observed (Fig. 2*a*). Binding of ecotin enlarges the active site relative to the benzamidine complex, pushing the two β-barrels apart. When applying the same

structural superposition to rat trypsin complexes with benzamidine (PDB code 1dpo; Earnest *et al.*, 1991) or ecotin (PDB code 1ezs; Gillmor *et al.*, 2000), a similar enlargement of the active site in the ecotin complex is observed relative to the benzamidine complex.

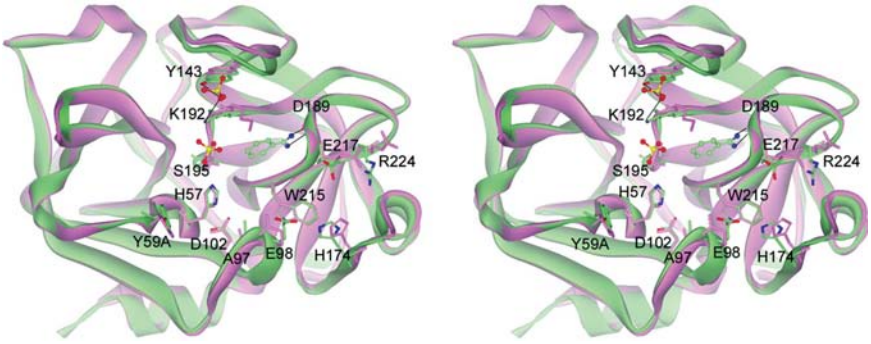

Figure 2
Stereoview of superposition of the rhFXI$_{370–607}$–ecotinM84R and rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine crystal structures. The two crystal structures are superimposed by aligning the C$^{\alpha}$ atoms of residues 31–122 and are presented as ribbon diagrams with rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine in green and rhFXI$_{370–607}$–ecotinM84R in pink. Several residues discussed in the text are shown in a stick representation, with N atoms in blue, O atoms in red and C atoms in green for the rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine structure and in pink for the rhFXI$_{370–607}$–ecotinM84R structure. The benzamidine and two sulfate molecules of the rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine structure are presented in a ball-and-stick representation. Hydrogen bonds are indicated by black lines.

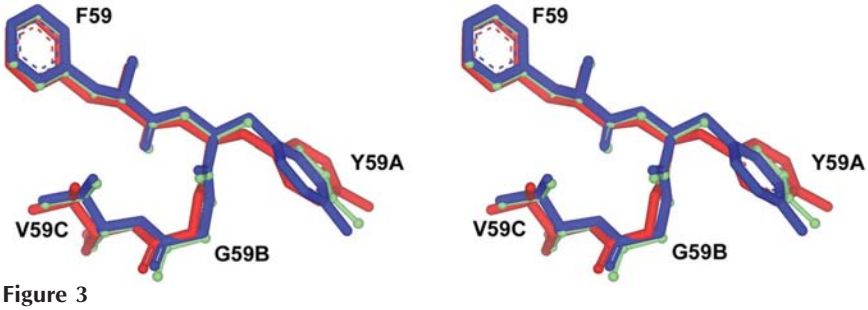

Figure 3
Superposition of *Tyr59A* in different rhFXI$_{370–607}$ structures. Residues 59–59*C* are presented. The rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine structure is in green, the rhFXI$_{370–607}$–ecotinM84R structure is in red and the structure of rhFXI$_{370–607}$-S434A, T475A,C482S,K437A in complex with an in-house compound is in blue.

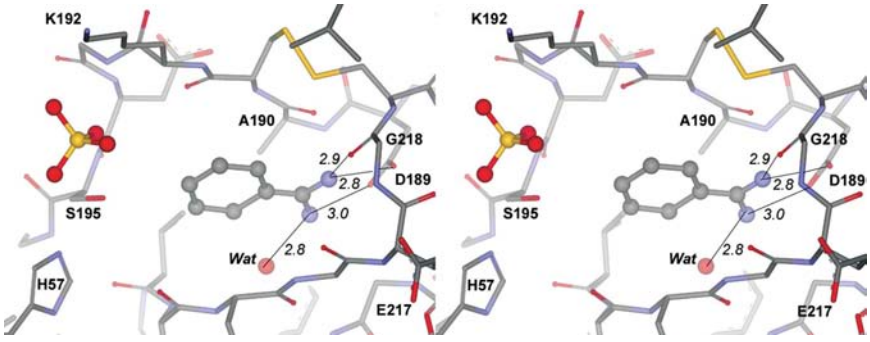

Figure 4
View of the benzamidine-binding site in the rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine structure. Benzamidine, a sulfate molecule in the oxyanion hole and a water molecule are shown as a ball-and-stick representation. The residues around benzamidine are shown as a stick representation. The hydrogen bonds between benzamidine and the protein are shown by black lines with bond distances labeled in Å.

Two unique features of rhFXI$_{370–607}$ in the rhFXI$_{370–607}$–ecotinM84R structure (Jin *et al.*, 2005) are also observed in the benzamidine complex structure. The conformation of the S_2 loop (residues *96–100*) that contains *Glu98* does not change significantly in the benzamidine-complex structure (Fig. 2). The loop folds into the active site with the aliphatic portion of *Glu98* packed against *Trp215*, thus decreasing the sizes of the S_2 and S_4 pockets in FXIa. The orientation of the carboxylic acid group of the *Glu98* is different in the two structures. In the benzamidine complex structure, one of the O atoms of *Glu98* forms a hydrogen bond with N$^{\varepsilon}$ of *His174* on the S_4 loop (residues *170–176*; Fig. 2). In the rhFXI$_{370–607}$–ecotinM84R structure, the *Glu98* side chain is oriented away from *His174*; one of the O atoms forms a hydrogen bond with its main-chain N atom and the other interacts with ecotin (Jin *et al.*, 2005). The C$^{\beta}$ and C$^{\gamma}$ methylenes of *Glu98* are in van der Waals interaction with Val81 of ecotin (P_4). In addition, Val81 of ecotin pushes the side chain of *Glu217* and *Arg224* away from the center of the active site, while these two residues face towards the center of the active site in the benzamidine complex structure (Fig. 2). *Tyr59A* in the two structures superimposes well (Fig. 2). It protrudes away from the surface of the protein and defines the unique shape of the S_2 pocket in FXIa. *Tyr59A* is in van der Waals contact with ecotin in the rhFXI$_{370–607}$–ecotinM84R structure but is packed against *Thr115*, *Asp116* and *Arg119* from a symmetry-related molecule in the rhFXI$_{370–607}$–benzamidine structure. We subsequently solved a structure of rhFXI$_{370–607}$ with another small-molecule inhibitor in space group $P2_1$ (data not shown). In this structure *Tyr59A* is in an orientation similar to the orientation in the ecotin and benzamidine complexes (Fig. 3), but it is not involved in crystal packing. This suggests that the conformation of *Tyr59A* is a true feature of rhFXI$_{370–607}$ rather than being induced by ecotin binding or crystal packing.

Lys192 is important for substrate and inhibitor binding (Babine & Bender, 1997). In the structure of the ecotinM84R complex, *Lys192* is disordered in one of the rhFXI$_{370–607}$ molecules (chain *A*) but completely ordered and interacting with ecotin in the other rhFXI$_{370–607}$ molecule (chain *B*). In the benzamidine complex, *Lys192* is ordered. A sulfate molecule,

trapped in the crystal from the crystallization solution, forms hydrogen bonds with *Lys192* and *Tyr143* (Fig. 2).

3.3. Benzamidine in the active site

Benzamidine is a weak inhibitor of FXIa. The IC_{50} of benzamidine against FXIa is around 100 μM in our fluorogenic assay. Benzamidine binds in the S_1 pocket with its amidine group hydrogen bonded to *Asp189* as well as the carbonyl oxygen of *Gly218* and a conserved water molecule (Fig. 4). This hydrogen-bonding network is the same as that seen with the guanidine group of Arg84 in one of the ecotinM84R molecules (chain *C*) in the rhFXI$_{370–607}$–ecotinM84R structure, as well as that observed for benzamidine in other trypsin-like serine proteases. There is a sulfate molecule in the oxyanion hole formed by O^{γ} of *Ser195* and two main-chain NH groups of *Ser195* and *Gly193*. This sulfate molecule has been observed at this location in many other serine proteases.

The benzamidine molecule in the rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine crystal can be easily replaced by other ligands. We have soaked different classes of FXIa inhibitors into the preformed rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine crystals. Fig. 5 shows three examples: a peptide mimetic with an α-ketothiazole moiety covalently attached to O^{γ} of *Ser195*, a noncova-

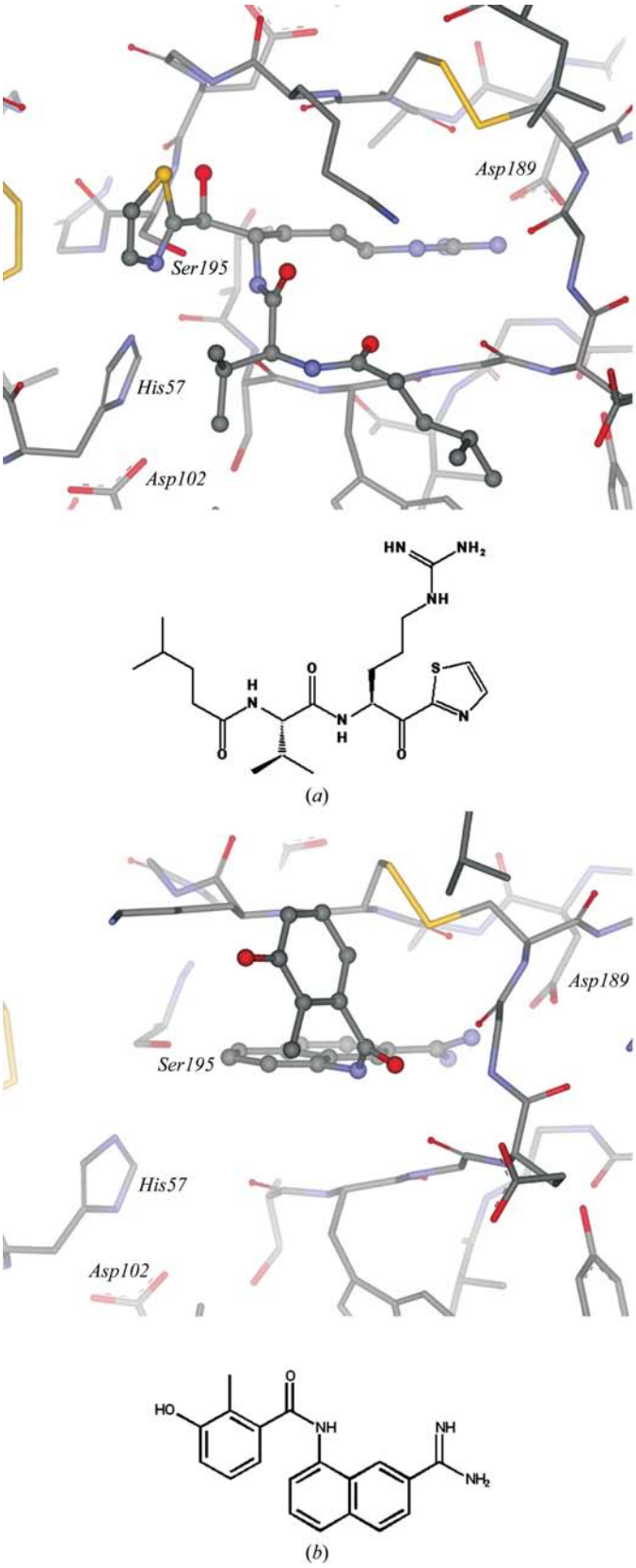

Figure 5
Examples of different classes of ligand soaked into rhFXI$_{370–607}$-S434A,T475A,C482S,K437A–benzamidine crystals. The active site of rhFXI$_{370–607}$-S434A,T475A,C482S,K437A is presented in a stick representation, with the catalytic triad (*Ser195*, *His57* and *Asp102*) and *Asp189* labeled. The ligands are in a ball-and-stick representation and their chemical structures are listed under each picture. (*a*) A peptide-mimetic ligand covalently connects to *Ser195*. (*b*) A naphthamidine ligand replaces benzamidine at the active site. (*c*) A non-basic small compound binds in the S_1 pocket.

lently bound naphthamidine compound and a non-basic fragment replacing benzamidine in the S_1 pocket. Our success with the mutant $rhFXI_{370–607}$ demonstrates that protein engineering is a powerful tool for improving protein crystallizability (Jin & Babine, 2004). Obtaining high-quality crystals is the key to iterative structure-based ligand design.

References

Babine, R. E. & Bender, S. L. (1997). *Chem. Rev.* **97**, 1359–1472.

Baglia, F. A., Gailani, D., Lopez, J. A. & Walsh, P. N. (2004). *J. Biol. Chem.* **279**, 45470–45476.

Baglia, F. A., Jameson, B. A. & Walsh, P. N. (1992). *J. Biol. Chem.* **267**, 4247–4252.

Baglia, F. A., Jameson, B. A. & Walsh, P. N. (1993). *J. Biol. Chem.* **268**, 3838–3844.

Baglia, F. A., Jameson, B. A. & Walsh, P. N. (1995). *J. Biol. Chem.* **270**, 6734–6740.

Baglia, F. A. & Walsh, P. N. (1996). *J. Biol. Chem.* **271**, 3652–3658.

Bouma, B. N. & Meijers, J. C. (2000). *Curr. Opin. Hematol.* **7**, 266–272.

Chung, C. H., Ives, H. E., Almeda, S. & Goldberg, A. L. (1983). *J. Biol. Chem.* **258**, 11032–11038.

Derewenda, Z. S. (2004). *Structure*, **12**, 529–535.

Earnest, T., Fauman, E., Craik, C. S. & Stroud, R. (1991). *Proteins*, **10**, 171–187.

Gillmor, S. A., Takeuchi, T., Yang, S. Q., Craik, C. S. & Fletterick, R. J. (2000). *J. Mol. Biol.* **299**, 993–1003.

Gruber, A. & Hanson, S. R. (2003). *Blood*, **102**, 953–955.

Hirsh, J., O'Donnell, M. & Weitz, J. I. (2005). *Blood*, **105**, 453–463.

Ho, D. H., Badellino, K., Baglia, F. A. & Walsh, P. N. (1998). *J. Biol. Chem.* **273**, 16382–16390.

Jin, L. & Babine, R. E. (2004). *Protein Crystallography in Drug Discovery*, edited by R. E. Babine & S. S. Abdel-Meguid, pp. 209–214. Weinheim: Wiley–VCH.

Jin, L., Pandey, P., Babine, R. E., Gorga, J. C., Seidl, K. J., Gelfand, E., Weaver, D. T., Abdel-Meguid, S. S. & Strickler, J. E. (2005). *J. Biol. Chem.* **280**, 4704–4712.

Longenecker, K. L., Garrard, S. M., Sheffield, P. J. & Derewenda, Z. S. (2001). *Acta Cryst.* D**57**, 679–688.

Longenecker, K. L., Read, P., Lin, S. K., Somlyo, A. P., Nakamoto, R. K. & Derewenda, Z. S. (2003). *Acta Cryst.* D**59**, 876–880.

Minnema, M. C., Friederich, P. W., Levi, M., von dem Borne, P. A., Mosnier, L. O., Meijers, J. C., Biemond, B. J., Hack, C. E., Bouma, B. N. & ten Cate, H. (1998). *J. Clin. Invest.* **101**, 10–14.

Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Ragni, M. V., Sinha, D., Seaman, F., Lewis, J. H., Spero, J. A. & Walsh, P. N. (1985). *Blood*, **65**, 719–724.

Walsh, P. N. (2001). *Thromb. Haemost.* **86**, 75–82.