PROTEINS: Structure, Function, and Genetics 30:264–274 (1998)

Mapping the Active Site of Factor Xa by Selective Inhibitors: An NMR and MD Study

F. Fraternali,[1] Q.-T. Do,[1] B.-T. Doan,[1] R.A. Atkinson,[1] P. Palmas,[1] V. Sklenar,[1] P. Safar,[3] P. Wildgoose,[2] P. Strop,[3] and V. Saudek[1]*

[1]*Marion Merrell Research Institute, HMR, Strasbourg, France*
[2]*Hoechst, Marion Roussel, Disease Group Cardiovascular, Frankfurt, Germany*
[3]*Selectide Corporation, HMR, Tucson, Arizona*

ABSTRACT **The structure of two selective inhibitors, Ac-Tyr-Ile-Arg-Ile-Pro-NH_2 and Ac-(4-Amino-Phe)-(Cyclohexyl-Gly)-Arg-NH_2, in the active site of the blood clotting enzyme factor Xa was determined by using transferred nuclear Overhauser effect nuclear magnetic resonance (NMR) spectroscopy. They represent a family of peptidic inhibitors obtained by the screening of a vast combinatorial library. Each structure was first calculated by using standard computational procedures (distance geometry, simulated annealing, energy minimization) and then further refined by systematic search of the conformation of the inhibitor docked in the active site and repeating the simulated annealing and energy minimization. The final structure was optimized by molecular dynamics simulations of the inhibitor-complex in water. The NMR restraints were kept throughout the refinement. The inhibitors assume a compact, very well defined conformation, embedded into the substrate binding site not in the same way as a substrate, blocking thus the catalysis. The model allows to explain the mode of action, affinity, and specificity of the peptides and to map the active site. Proteins 30:264–274, 1998.** © 1998 Wiley-Liss, Inc.

Key words: factor Xa; serine proteinases; blood coagulation; active site inhibitors; transferred NOE

INTRODUCTION

Blood coagulation enzymes are serine proteinases highly specialized in the fine tuning of activation, maintenance, and regulation of blood clot formation (e.g., factors VIIa, IXa, Xa, thrombin) and clearance (e.g., plasmin). Their catalytic domains are homologous in the amino acid sequence and share a common overall three-dimensional chymotrypsinlike fold.[1] Their primary trypsin-type specificity is also common, as they all hydrolyze peptide bonds following an arginine. Despite this global similarity, they have evolved remarkably specialized substrate specificity.[2] Whereas the trypsin-type specificity, conferred by the binding pocket S1, is well understood,[1] the details of the further binding sites have been addressed only recently, employing methods such as X-ray crystallography[3–6] site-directed mutagenesis,[7] and modeling.[8,9] An alternative approach is to map the binding sites by identifying selective inhibitors. A detailed understanding of the specificity is important not only from the theoretical point of view but also therapeutically, since the blood-clotting enzymes are attractive targets for treatment of thrombosis.

Factor Xa, a typical 45-kD serine proteinase, is located at a crossroads of the coagulation cascade[10]: it is activated from factor X through limited proteolysis by either factor VIIa (extrinsic initiation pathway) or IXa (intrinsic maintenance pathway). Its main function is to activate thrombin, but it is also autoactivating and participates in activation of factor VIII. Its selective inhibition would block both pathways. Indeed, the use of tick anticoagulant peptide,[11] a 60-amino acid protein inhibiting selectively factor Xa, has demonstrated the feasibility of this approach to thrombosis.[12]

Recently, a family of tri- to pentapeptides acting as competitive inhibitors of FXa has been identified[13] by screening millions of peptides in a combinatorial library.[14] Their consensus sequence is Tyr-Ile-Arg-Ile-X, where X is a hydrophobic amino acid and Ile may be replaced by Leu. The minimal sequence consists of the first three residues, which may be also substituted by homologous nonnatural amino acids.

Abbreviations: One or three-letter standard abbreviations were used for the amino acids; DG, distance geometry; EM, energy minimization; FXa, factor Xa; MD, molecular dynamics; RMSD, root-mean-square deviation; SA, simulated annealing; 3D, three-dimensional;

Present addresses: F. Fraternali, EMBL D-6900 Heidelberg, Germany; Q.-T. Do, Tripos Associates, F-92167 Antony, France; B.-T. Doan, LCOSB, Universite Pierre et Marie Curie, F-75252 Paris, France; R.A. Atkinson, UPR 9003, ESBS 67400 Illkirch, France; P. Palmas, CEA, F-37260 Monts, France; P. Safar, Masaryk University CZ-611 37, Brno, Czech Republic; V. Saudek, Synthélabo Biomoléculaire, F-67080 Strasbourg, France

*Correspondence to: Dr. Vladimir Saudek, Synthélabo Biomoléculaire, F-67080 Strasbourg, 16 rue d'Ankara, France.
E-mail: saudekv@synbio.tpgnet.net

Received 20 March 1997; Accepted 1 August 1997

TABLE I. Chemical Structure and Specific Activity of Factor Xa Inhibitors

Peptide	Amino acid sequence*	K_i (μM)†
A	Ac-Tyr-Ile-Arg-Ile-Pro-NH_2	1.6
B	Ac-(4-Amino-Phe)-(Cyclohexyl-Gly)-Arg-NH_2	0.3

*All amino acids are L-isomers, the three letter code is used for the standard amino acids and the sequence is numbered from 0 for acetyl group to 6 for the NH_2 group for peptide A and similarly from 0 to 4 for peptide B.
†Competitive inhibition with the substrate Carbobenzoxy-D-Arg-L-Arg-4-naphtyl anilide.[13]

They are highly selective, with at least two orders of magnitude greater potency toward FXa than to other trypsinlike serine proteinases such as factor VIIa, thrombin, trypsin, plasmin, proteinase C, kallikrein, and tissue plasminogen activator.[13]

To better understand the mode of action and selectivity of these inhibitors and the active site of FXa, we undertook the study of their active conformation. We selected two members of the family (Table I), the first being a typical member, the second corresponding for the shorter inhibitors built from nonnatural amino acids. Although the crystal structure of FXa has become recently available,[5] attempts to obtain crystals of the FXa complex with these inhibitors in the active site have been unsuccessful to date (H. Schreuder, personal communication, 1996). On the other hand, NMR spectroscopy may be applied directly to study low molecular weight ligands in interaction with their macromolecular receptors by measuring transferred NOEs, providing the ligands are in fast exchange between their receptor-bound and free state.[15,16] Here we describe the three-dimensional structure of the selected peptides in the active site and discuss their mode of inhibition and interaction with the active site.

MATERIALS AND METHODS

Sample Preparation

Human FXa (E.C. 3.4.21.6) was purchased from Enzyme Research Laboratories Inc. It was found to be homogeneous on 10% SDS-PAGE gels and fully active toward the chromogenic substrate Cbz-D-Arg-Gly-Arg-paranitroanilide. The enzyme was initially solubilized in a 20 mM Tris-HCl buffer pH = 7.4 containing 0.7 M NaCl and stored at −80°C. It was washed three times with a 50 mM phosphate buffer pH = 7.0 at 4°C by ultracentrifugation employing Filtron Centricon filters with a molecular weight cutoff of 10 kD, to remove all traces of low molecular-weight impurities, and finally concentrated to about 0.1 mM. The exact concentration was then established by measuring UV absorbance at 280 nm.

The inhibitory peptides were prepared by solid state synthesis and purified by reverse phase HPLC as described[13] (Selectide US Patent, PTC/US95/05268). Their chemical structure and activity is given in Table I. The inhibition constants K_i were measured as follows[13]: FXa (0.5 nM) was preincubated for 15 minutes with an inhibitor, a chromogenic substrate Cbz-(D)Arg-Gly-(L)Arg-*p*-nitroanilide (Pharmacia) was added to a final concentration of 175 μM and the release of nitroaniline was monitored. K_i was calculated from the slope of the linear part of the progress curves registered with 8 different inhibitor concentrations. The samples for NMR measurement contained 2 mM inhibitor and 0.02–0.1 mM FXa (inhibitor:enzyme molar ratio 20:1, 40:1 and 100:1) in 0.5 ml of 50 mM phosphate buffer pH = 7.0, 90% H_2O, 10% D_2O in 5-mm NMR tubes. To prepare samples in D_2O, the enzyme was Centriconed three times in the buffer prepared in D_2O, and the peptides were lyophilized from solutions in D_2O.

NMR Methods

NMR experiments were carried out on Bruker AMX-500 and DMX-600 MHz spectrometers. A full range of 2D NMR spectra (COSY, TOCSY, ROESY, NOESY) of the peptides in the absence and the presence of the enzyme were recorded at 37, 24, and 10°C. All spectra were acquired in the States-TPPI mode.[17] In order to prevent small temperature variations of proton chemical shifts in different experiments, a CITY pulse train[18] was introduced before each sequence except TOCSY to preheat the sample. Since factor Xa is active at neutral physiological pH, water suppression applying saturation was avoided and several variants of the WATERGATE sequence[19,20] were applied instead. To reduce water signal broadening in ω_1 dimension, radiation damping was attenuated by weak bipolar gradients.[21] This allowed all HN resonances to be detected unperturbed, with high-intensity exchange cross-peaks with H_2O in the 2D NOESY, TOCSY, and ROESY spectra. No dipolar interaction between H_2O and the ligands was observed. The COSY spectra were recorded with a sweep width of 12.1 ppm, 32 accumulations, with 512 data points in ω_1 and 2048 in ω_2; the TOCSY[18] spectra with a mixing time of 60 ms, a spin-lock field of γB_1 = 11.4 kHz, with cross-relaxation compensation and *z* purging field gradients, 24 accumulations, 256 data points in ω_1 and 2048 in ω_2. The ROESY spectra were acquired with 32 accumulations, a 100 ms spin-lock field of γB_1 = 5.6 kHz, with 320 data points in ω_1 and 2,048 in ω_2, the NOESY spectra with homospoil gradients during the mixing time to suppress radiation damping during this period, 32 accumulations, 320 data points in ω_1 and 2048 in ω_2. Off-resonance ROESY experiments[22] were performed with an adiabatic trapezoidal spin lock pulse of 1.4 ms increasing time, a 8 kHz spin-lock field, a 100 ms mixing time, 6 offsets corresponding to angles θ = 54.7°, 50.8°, 47.0°, 43.5), 40.0°, and 35.3° between the rotating frame and the static magnetic field. The 1.4-ms pulse was cali-

brated in such a way that it induced no significant modification of the water 1D spectrum. A function of $\cos^2 \theta$ versus peak volume was analyzed in terms of σ_{ROE} and σ_{NOE} cross-relaxation rates. The duration of the mixing times in ROESY and NOESY was randomly varied ($\pm 0.5\%$) in order to cancel magnetisation due to zero-quantum and single-quantum coherences.

The spectra were processed and analyzed using UXNMR and AURELIA software packages (Bruker Spectrospin). The assignments were obtained using the standard procedures.[23] The NOE buildup of the 2D transferred NOEs spectra[16] recorded with the mixing times of 50, 100, 150, 200, 250, 300, 400, and 500 ms was found to be linear at up to mixing time of 200 ms and provided interproton distances. The values for fixed distances in methylene groups were used for calibration. Further upper-distance limits defined for 50%, 25%, 12.5%, and so on of the methylene volume, using an r^{-6} relationship, were used up to 6.0 Å. Pseudoatoms corrections[23] were applied when necessary.

Computational Methods

All calculations were performed with the molecular modeling software package SYBYL 6.2 (Tripos Associates, St. Louis, MO, USA) on a Silicon Graphics Indigo 2 Extreme work station. The Tripos force field[25] was utilized without the electrostatic term for initial refinements where the presence of the enzyme was neglected. The Kollman all-atom force field[25] with a dielectric constant of 1 and a cutoff of 8 Å for the nonbonded interactions was applied in the calculations with the enzyme. DG calculations were performed with the DIANA program[26] as implemented in SYBYL. The nonstandard amino acids were added to the DIANA and SYBYL libraries. Structures were generated starting from conformations with randomly distributed torsion angles and using all nonredundant intra- and interresidual distance constraints compiled from the NMR spectra. EM consisted of 200 steps of simplex optimization procedure followed by 5,000 iterations with the Powell method[27,28] with a gradient tolerance of 10^{-3} kcal $\cdot$ mol$^{-1} \cdot$ Å^{-1}. The SHAKE[29] algorithm was applied to all bonds involving hydrogen atoms. SA was carried out as implemented in SYBYL (100 cycles consisting of heating to 1,000 K, equilibration at 1,000 K for 5 ps and linear cooling from 1,000 K to 0 K for 10 ps) for the initial refinements. When the enzyme was kept as a rigid charged template around the inhibitor, the following modifications were introduced in the SA protocol: the temperature was raised from 0 to 1,000 K in 10 steps of 0.5 ps, maintained at 1,000 K for 20 ps, and lowered to 0 K in 10 steps of 0.5 ps. The systematic search examined energies of structures generated by rotating selected rotatable bonds.

MD simulations were performed at 300 K in explicit boxes of SPC (single-point charge) water. The Silverware algorithm[30] was employed to construct a water cage around the system[31] at constant volume. Molecular accessible surface areas were calculated with WHATIF.[33] The values are reported relative to those of the exposed residue in the tripeptide Ala-X-Ala, where X is the residue of interest.

The structure determination followed the flow chart in Figure 1. The NMR distance range constraints were retained throughout the whole computational procedure. A force constant of 35, 10, and 5 kcal mol^{-1} Å^{-2} was used in SA, systematic search, and EM, respectively. One thousand structures were generated in the distance geometry calculations, and structures with the lowest target function value were submitted to SA for refinement. Among the resulting 100 energy minimized structures, the lowest energy structure was selected for further refinement. The guanidine group of Arg3 was superimposed on the guanidine group of Arg51′ (from the neighboring molecule in the crystal lattice) in the active site of the crystal structure of FXa. A systematic conformational search was performed on the rotatable bonds C^{α}—C^{β}, C^{β}—C^{γ}, C^{γ}—C^{δ} and C^{γ}—C^{ϵ} of Arg3 of the peptide docked in the active site. Standard SYBYL parameters for bump check (general, 1–4 and hydrogen bond van der Waals scale factors 0.95, 0.87, and 0.65, respectively) were used. The enzyme was kept as a rigid template and the angles were incremented in 5-degree steps. The accepted resulting structures were anchored in the enzyme by defining two distance range constraints of 3 to 5 Å (force constant of 10 kcal $\cdot$ mol^{-1} $\cdot$ Å^{-2}) between the NH of Arg3 guanidine and the carboxyl oxygens of Asp189 in FXa and resubmitted to SA and EM. The lowest energy conformers were optimized by MD and EM in water. In these final simulations both the inhibitor and the enzyme were allowed to move, in order to relax possible artificial strains induced by the enzyme.

The sequence alignment of the selected serine proteinases was done using the GCG Wisconsin Package, Version 8, with Fasta, Blast, and PileUp routines (Genetic Computer group Manual, 1994). The enzyme sequences were extracted from the Swiss Protein Database. The crystallographic structure of FXa[5] was taken from the Brookhaven database, entry 1hcg. The residues are numbered according to the chymotrypsinogen sequence, as suggested by Bode and colleagues.[6]

RESULTS AND DISCUSSION

NMR Spectra

An example of the NMR spectra is shown in Figure 2. Both peptides A and B (Table I) provided clear TOCSY and COSY spectra allowing the NMR assignment of all amino acid types. Almost no cross-peaks

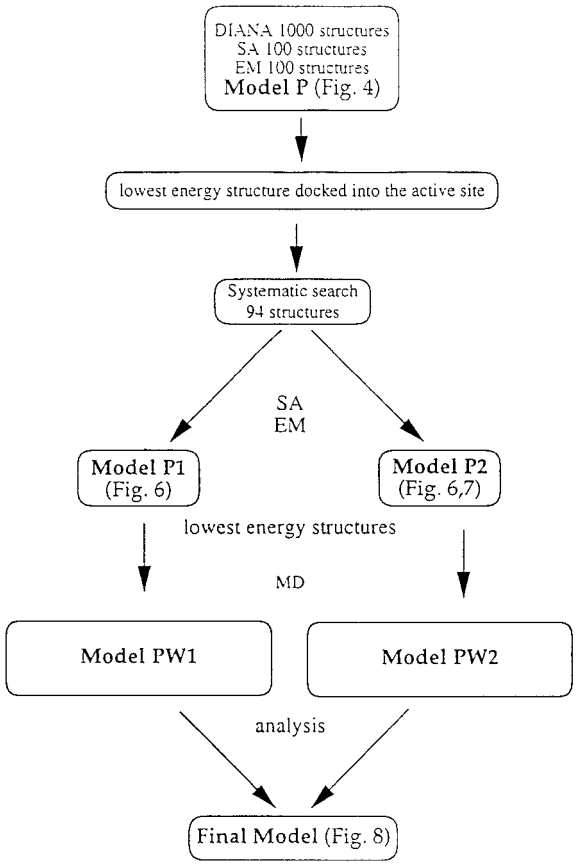

Fig. 1. Flow chart of the computational procedures applied for the structure determination and refinement.

appeared in the NOESY spectra (Fig. 2b). ROESY spectra (Fig. 2c) showed only a limited number of intraresidue and sequential NOEs sufficient to complete the sequential assignment (Table II). Simultaneous measurements of σ_{NOE} and σ_{ROE} from a series of off-resonance ROESY gave internuclear (14 intraresidual and 7 sequential nonredundant) (Fig. 3) distances. The 10 best distance geometry structures of peptide A, refined by SA and EM displayed no distance violations greater than 0.05 Å. Their conformation appears as two successive loose turns in opposite directions (Fig. 4a); however, the average conformation is extended. The mean RMSD value of 4.1 Å for their superposition for all atoms and of 2.4 Å for the backbone atoms indicates a poor definition of the overall conformation. Very similar results were obtained for peptide B. Clearly, the peptides free in solution do not adopt any specific structure.

No chemical shift changes were found in the spectra of the peptides after addition of factor Xa at molar ratios varying from 0.01 to 0.05. The NOESY spectra changed however remarkably, displaying a great number of well-resolved negative transferred NOEs between the peptide protons (Fig. 2a,c). Few additional weak NOEs arising probably from the flexible parts of the protein were also observed. These could be distinguished from the cross-peaks of the peptides given the assignments of the peptide resonances and by comparison with the spectra of the protein alone. When the temperature of the NOESY experiments was varied between 5 and 40°C, the NOEs become more intense at higher temperature. The spectra did not allow any clear NOE between the protein and the peptide to be discerned. Only positive transferred ROE cross-peaks were observed in the ROESY spectra, indicating that no indirect, protein mediated intermolecular NOEs were present.[34] A control spectrum acquired under the same conditions as the NOESY spectra, but with the factor Xa replaced by serum albumin, showed that the transferred NOE effect was not due to an increase in viscosity. The transferred NOEs were also completely suppressed in a sample where 0.01 mM dansyl-Glu-Gly-Arg-CH_2Cl was added 24 hours prior to the addition of the inhibitors under study. This specific covalent inhibitor blocs the active site,[13] showing that the observed transferred NOEs must be due to the specific interaction of the peptide in the active site of the enzyme. The occurrence and the temperature dependence of the transferred NOEs indicates that the inhibitors are in the fast exchange between their enzyme-bound and free state.

The best signal-to-noise ratio was obtained in the spectra acquired at 20°C and enzyme–peptide molar ratio of 0.01. The structures of the peptides were therefore determined under these conditions. NOESY spectra of the ligand in the absence of the enzyme contained no cross-peaks, consequently, no correction for the NOEs cross-peaks of the free peptide was necessary. No spin diffusion was observed in the NOE buildup curves up to mixing times of 200 ms; 14 intraresidue and 54 interresidue nonredundant distances were measured for peptide A, while 7 and 16 distances for the peptide B. The average number of NOEs per detected proton resonance in the individual amino acids is summarized in Table II for peptide A and in Table III for peptide B. The distribution of the interresidue NOEs over the sequence is illustrated in Figure 3 for peptide A. Such a high number of NOEs is usually observed only for residues buried within the core of a protein. It is in sharp contrast with the low number of NOEs detected in the peptides free in solution. This qualitative inspection of the NOESY spectra indicates that the enzyme forces the peptides to adopt a compact conformation.

The spectra of peptide A in the presence of FXa remained unchanged for at least 2 months, indicating that the peptide is resistant to any proteolytic action by FXa. A second set of resonances with slightly different chemical shifts was detected in the spectra of peptide B after 3 weeks in the presence of FXa. The greatest chemical shift change was apparent on the Arg3 cross-peaks. The most likely explanation of the progressive change in intensity of the

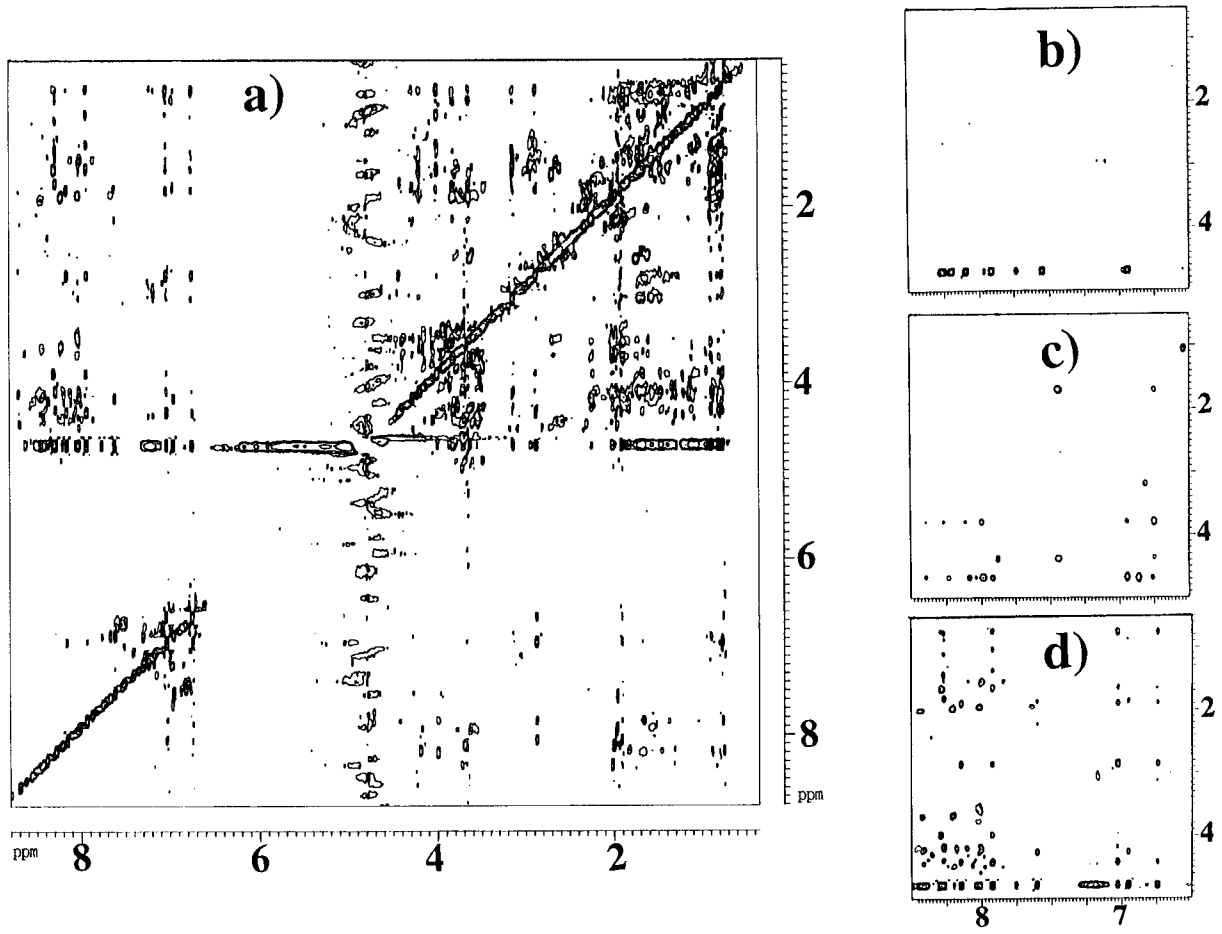

Fig. 2. 600 MHz NMR spectra of 2 mM peptide A in 0.05 M phosphate buffer pH 7.40 at 20°C. **a** and **d:** NOESY spectrum of the sample in the presence of 0.02 mM factor Xa (mixing time 100 ms). **b:** NOESY (mixing time 300 ms). **c:** Off-resonance ROESY (mixing time 100 ms) spectra, sample in the absence of factor Xa.

TABLE II. ^{1}H NMR Assignment of Peptide A at 20°C, pH 7.40

Residue	NH	Hα	Hβ	Others	⟨NOE⟩*
Ac0				CH_3 1.90	4.0
Tyr1	8.10	4.54	2.96	2,6H 7.13	6.4
				3,5H 6.84	
Ile2	7.92	4.11	1.77	γ CH_2 1.14	6.0
				γ CH_3 1.46	
				δ CH_3 0.85	
Arg3	8.26	4.31	1.73/1.80	γ CH_2 1.56/1.63	2.4
				δ CH_2 3.77	
Ile4	8.20	4.47	1.91	γ CH_2 1.19/1.52	4.3
				γ CH_3 0.97	
				δ CH_3 0.87	
Pro5		4.36	1.97/2.31	γ CH_2 2.08	2.8
				δ CH_2 3.71/3.87	
(NH_2)6				NH_2 6.94/7.56	—

*Average number of NOEs per detected ^{1}H resonance.

second form is a very slow hydrolysis of the peptide bond Arg3-(NH_2)4. The pattern of the NOESY peaks was identical in the newly emerging set of peaks, indicating that the resulting fragment retained affinity toward FXa and adopted a similar conformation as the intact peptide.

3D Structure of the Inhibitors

The structures with the lowest total energy resulting from the distance geometry calculations followed by refinement by SA and EM (model P) are overlaid for minimal RMSD for peptide A in Figure 4b and the structures for peptide B in Figure 4c. Their structural statistics are given in Table IV. All the NMR constraints were well satisfied (the largest distance violation was 0.44 Å) and compatible with a single conformational family within the limits of the experimental error. There is no significant deviation from the standard geometries. The conformation of both peptides is very well defined not only for the backbone but also for the side chains as judged according to the global RMSDs in Table IV and by inspection of the superpositions in Figure 4. Similarly, the local RMSD of the individual backbone heavy atoms are distributed uniformly and, with the exception of the initial acetyl and terminal NH_2 groups are always

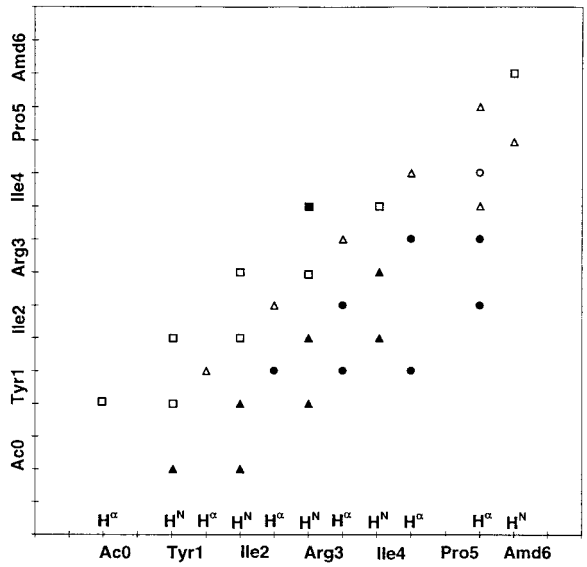

Fig. 3. Summary of the NOE cross-peaks detected in the off-resonance ROESY spectra of the free peptide A (*open symbols*) and in the TrNOE spectra of the peptide A in the presence of factor Xa. All ROESY cross-peaks were also detected in the NOESY spectra, backbone–backbone, *filled square;* backbone–side chain *filled triangle;* side chain-side chain connectivities *filled circle.* Backbone NH and CαH protons are given separately for each residue.

lower than 0.5 Å (Fig. 5, curve P). A high degree of the global and local structure definition is never achieved with low molecular weight peptides free in solution. Clearly, the inhibitors are located inside the active site of the enzyme and their motion must be restrained by the surrounding protein. The interaction of the inhibitor with the enzyme induces a very compact and well-defined structure.

The only exception to this tight definition of the structure is the side chain of the arginine, especially the atoms following the $CH_2\beta$ group. Overlaying the structures for this side chain alone leads to a rise of the local RMSD (Fig. 5b). Inspection of the 3D structures in Figure 4b,c shows that the side chain of Arg 3 always protrudes from the compactly folded body of the peptides and shows no contact with the remaining side chains nor with the backbone. It cannot therefore give any dipolar coupling with them.

Position of the Inhibitors in the Enzyme

In the standard computational procedures of NMR structure determination (model P, first stage in Fig. 1), the presence of the enzyme atoms in the vicinity of the inhibitor is not taken into consideration and the molecule is treated by the force field as if it were in vacuo. However, the inhibitor is known to compete with the substrate for the active site (Table I) and can be displaced by transition state inhibitors that block the active site.[13] The peptides must therefore be located in the active site of the enzyme. We can add this information to the NMR results and further refine the structure. By retaining all NMR constraints, we can also apply the experimental information for positioning the inhibitors in the active site. This approach is based on the assumption that there is no major conformational change in the enzyme upon binding the inhibitor. However, a major conformational change is unlikely, as we start from a crystal structure, the active site of which is occupied by a natural ligand—arginine. On the other hand, a small adjustment of the side chains is allowed in the final stages of the refinement using MD. To avoid any subjective bias, we did not use manual docking but added instead a systematic search to the standard NMR methods (SA and EM, Fig. 1).

The primary site of the interaction of a substrate or a competitive inhibitor with a trypsinlike serine proteinase is the substrate pocket S1. The substrate/inhibitor arginine (or lysine) side chain is embedded in a hydrophobic cavity forming a salt bridge through its positively charged end group with Asp189 at the bottom.[1] Clearly, for this very reason no NOEs between the side chain protons of the arginine and the rest of the inhibitors could be detected. The S1 pocket isolates them from all contacts with the rest of the peptide. We have therefore obtained so far a well-defined body of the inhibitors (Fig. 4b,c), suitable for docking, but the anchoring moiety of Arg3 remained underdetermined. One crystallographic unit cell of FXa contains 8 molecules and the C-terminal arginine of the A-chain (Arg51′) of one molecule fills the S1 specificity pocket of the neighboring molecule in a substratelike manner. The position of the guanidine group can serve as a starting point for a systematic search of acceptable locations of the inhibitor. If we rotate all the bonds in the side chain of Arg3, the conformational space available for the inhibitor inside the active site is systematically explored. Thus 94 energetically acceptable conformations, divided into 2 families of similar structures, were obtained. These were used as starting structures for further refinement by SA and EM, this time keeping the peptide inside the enzyme. The global values of energy of the resulting structures and the range and number of violations (data not shown) were essentially identical to those in Table IV obtained for the model P (Fig. 1) at the first stage of structure calculation. This indicates that the additional constraint imposed in the calculation by the enzyme does not deteriorate the quality of the structure. In other words, the experimental NMR constraints are fully compatible with the force field describing the interior of the active site. The global RMSDs were also similar (1.12 Å for the backbone, 1.63 Å for all heavy atoms [cf. Table IV]), but this time the arginine side chain could be included into the fit. Figure 6 shows 14 low-energy structures. The division into two families, observed already after the systematic search, persisted in the resulting structures, henceforth labeled as models P1 and P2 (Figs. 1 and 6), although individual structures could move from one family to the other during SA. Clearly, the

TABLE III. ^{1}H NMR Assignment of Peptide B at 20°C, pH 7.40

Residue	NH	Hα	Hβ	Others	⟨NOE⟩*
Ac0				CH_3 1.90	2.0
(4-Amino-Phe)1	8.13	4.42	2.82/2.87	2,6H 6.97	2.8
				3,5H 6.70	
(Cyclohexyl-Gly)2	8.02	4.03	1.62	γ CH_{2A} 0.88/0.94	1.7
				γ CH_{2E} 1.61/1.45	
				δ CH_{2A} 1.13/1.15	
				δ CH_{2E} 1.63/1.65	
Arg3	8.31	4.17	1.69/1.78	γ CH_2 1.56/1.58	0.5
				δ CH_2 3.14	
(NH_2)4				NH_2 7.07/7.56	—

*Average number of NOEs per detected ^{1}H resonance.

TABLE IV. Structural Statistics

Contributions*	Peptide A	Peptide B
Bond stretching	2.8	2.1
Angle bending	18.1	6.2
Torsion	15.0	7.3
Out of plane	0.3	0.1
Van der Waals 1–4	8.6	6.9
Van der Waals	−18.3	9.2
Fixed distance	5.4	1.4
Total	31.9	16.0
$\langle RMSD_{tot} \rangle^{\dagger}$	1.49	1.37
$\langle RMSD_{bb} \rangle^{\ddagger}$	1.12	1.56
Violations§		
>0.4 Å	1	0
>0.3 Å	1	1
>0.2 Å	4	3
>0.1 Å	3	4

*Average energy (kcal · mol^{-1}) and RMSD (Å) of the structures obtained after DG, SA, and EM (model P in Figs. 1 and 4).
†All heavy atoms except the side chain of Arg3.
‡Backbone heavy atoms.
§Number of violations exceeding the specified distance.

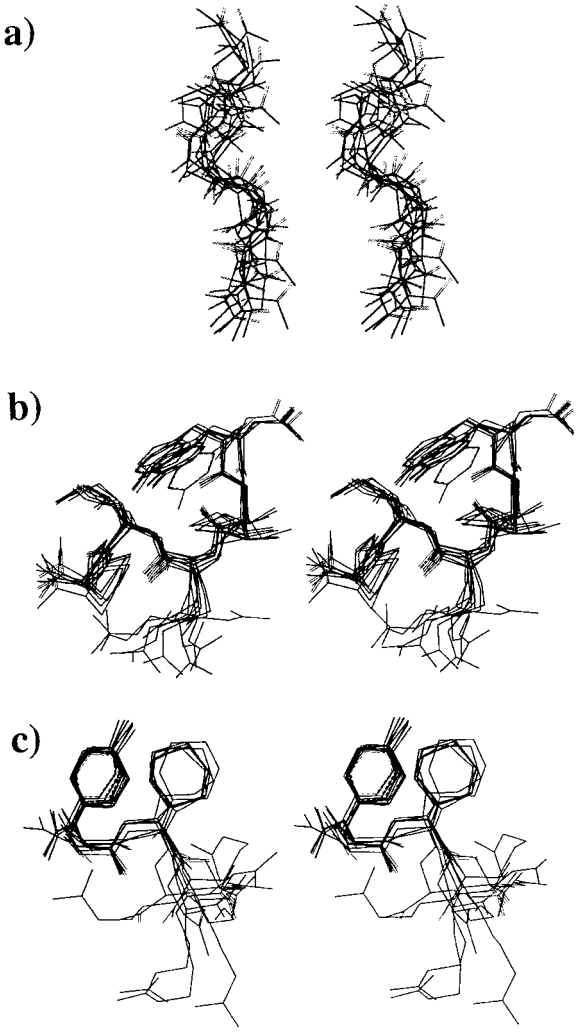

Fig. 4. Stereoview of the lowest energy distance geometry structures refined by SA and EM and overlaid for the minimal pairwise RMSD values of the backbone heavy atoms. **a:** Peptide A free in solution (C terminus is at the top). **b:** Peptide A in the presence of FXa. **c:** Peptide B in the presence of FXa.

systematic search as well as the following SA identified two minima of similar energy. The definition of the structure within each model improved significantly as compared to model P (cf. Figs. 4b and 6). The backbone RMSD (cf. Table IV) dropped to 0.70 Å and 0.71 Å, the RMSD for all atoms including the arginine to 1.22 Å and 0.98 Å for models P1 and P2, respectively. Figure 5a compares the improvement of the backbone RMSD along the chain. This is particularly apparent on the relative position of the arginine side chain with respect to the main body of the inhibitor (Fig. 5b). Both models, when overlaid with the structures P, were contained within the set of structures for model P.

The principal difference between the models P1 and P2 consists in the positioning of the inhibitor in the active site. Model P2 is shown in Figure 7. Table V compares the fractional surface areas for the individual residues in the peptide. In particular, Ile2 of P2 is embedded in the hydrophobic region formed by Tyr 99, Phe174, and Trp215 (referred to as substrate pocket S4[5]). Model P1 places Ile4 in this pocket. P1 and P2 are separated by a large energy barrier, such that to obtain a transition from one state to the other, the temperature had to be raised

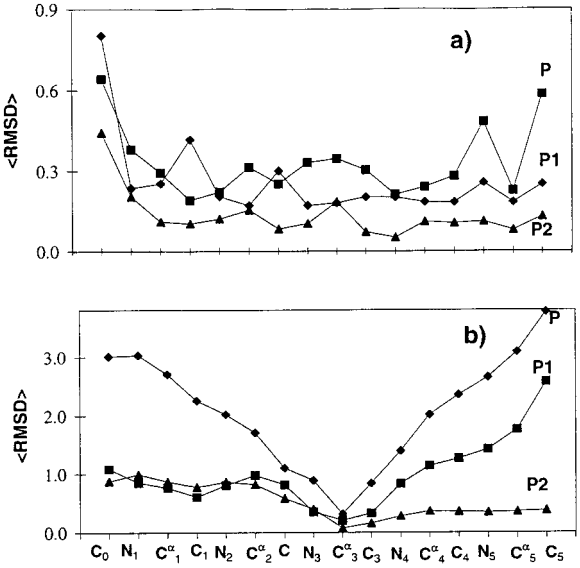

Fig. 5. Average pairwise RMSD(Å) backbone CO, C^{α} and N atoms of peptide A. P, P1 and P2 refer to the structures in Figure 1. **a:** All backbone. **b:** Only the Arg3 side chain atoms were superimposed.

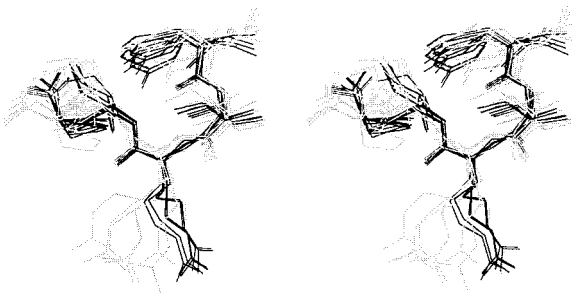

Fig. 6. Stereoview of 14 structures with the lowest energy obtained after systematic search, SA and EM, where the presence of FXa was taken into consideration. P1 structures are in gray, P2 in black, they are overlaid for minimal RMSD of all heavy atoms.

TABLE V. Comparison of Molecular Accessible Areas of the Individual Amino Acids in Model PW1 and PW2

Amino acid	PW1	PW2
Tyr1	33.31	37.01
Ile2	7.77	5.53
Arg3	4.69	3.13
Ile4	33.67	27.09
Pro5	16.90	4.11

TABLE VI. Energy Contributions (kcal/mol) to the Formation of the Inhibitor–Enzyme Complex

Model in vacuo*	P1	P2
Enzyme	1535 ± 0	1534 ± 0
Peptide	39 ± 3	36 ± 6
Complex	1475 ± 16	1448 ± 3
Peptide–enzyme	−100 ± 16	−123 ± 4
Model in water*	**PW1**	**PW2**
Enzyme	31	23
Peptide	15	20
Complex	−45	−89
Peptide–enzyme	−91	−132
Peptide–solvent	−139	−116
Enzyme–solvent	−1433	−1123
Complex–solvent	−1663	−1372
Solvent–solvent	−112	−486

*Energy contributions of the atoms within 10 Å radius from the Cα of the inhibitor Arg3 are given. Enzyme, peptide, complex stand for the enzyme energy, peptide energy, total energy contributions respectively; peptide–solvent, enzyme–solvent, complex–solvent, solvent–solvent represent the respective energy of interaction. Average energies (±standard deviation) are given for 10 lowest energy structures in each model P1 and P2.

to 1,000 K in SA, while the two conformers remained separated in a run with a 500 K maximum. Clearly, since we cannot consider that the two conformers could be exchanging in reality, only one model has to be selected. The 3D structure of the inhibitor in model P2 is better adapted to the active site: the global RMSD is lower (1.22 vs 0.98 Å), its distribution along the backbone goes through smaller values (Fig. 5a) and the position of the C-terminal fragment is much better determined (Fig. 5b). Table VI, where the average energy of 10 structures in each family is broken into various contributions shows that, for the simulations performed in vacuo, the intrinsic peptide energy is slightly lower in model P2, but more importantly, the intracomplex energy and inhibitor-enzyme energy are, respectively, 27 and 23 kcal · mol^{-1} in favor of P2.

To evaluate the energy of the two models in a force field as complete as possible, a 65 ps MD run in water was performed on the lowest energy representative of each model. The structure of the enzyme did not change significantly during MD: for example, the RMSD between the initial structure and the structure in model PW2 is 1.39 Å for all atoms and 0.98 Å for backbone atoms. Table VI reports the energy of the resulting structures after EM (models PW1 and PW2). Similarly, as at the preceding stages of the refinement, the initial fixed distance energy and the number and distribution of the violations given in Table IV remained unchanged.

We observe that for the simulations in vacuo, all contributions to the inhibitor-enzyme complex formation (complex and interaction contributions) are in favor of model P2. However those results are not conclusive because the enzyme is still considered as a rigid template at this stage. More detailed analysis of the two models can be outlined from the simulations in water. The interaction energy contribution is also in these simulations more favourable for PW2, whereas the interaction energy with the solvent is

TABLE VII. Alignment of the Amino Acid Sequences in the Vicinity of the Substrate Site S4 in Selected Human Serine Proteinases*

		99†	174	215
Factor Xa	HNRFT.K.....	ET**Y**DFDIAV	SSS**F**IITQNM	VTGIVS**W**G.EGCA
Factor VIIa	PSTYV.P.....	GT**T**NHDIAL	GDS**P**NITEYM	LTGIVS**W**G.QGCA
Proteinase C	HPNYS.K.....	ST**T**DNDIAL	VMS**N**MVSENM	LVGLVS**W**G.EGCG
Thrombin	HPRYNWR.....	EN**L**DRDIAL	STR**I**RITDNM	QMGIVS**W**G.EGCD
Kallikrein	HPLYNMSLLKHQSLRPD	ED**S**SHDLML	AYS**E**KVTEFM	LQGITS**W**GPEPCA
Trypsin	HPQYDRK.....	.T**L**NNDIML	SYP**G**KITSNM	LQGVVS**W**G.DGCA
Plasmin	EP.....	..**T**RKDIAL	FLN**G**RVQSTE	LQGVTS**W**GL.GCA
TPA‡	HKEFDDD.....	DT**Y**DNDIAL	LLN**R**TVTDNM	LVGIIS**W**GL.GCG

*One letter amino acid code, the bold letters indicate the residues which were found in close contact with the hydrophobic side chain at position 2 in the inhibitors A and B in factor Xa and their counterparts in the other enzymes.
†Chymotrypsin sequence numbering.
‡Tissue plasminogen activator.

more favorable for PW1. The inhibitor in PW1 exhibits a larger number of hydrogen bonds with water (10 in total) and less with the enzyme (2). In contrast, model PW2 shows 7 hydrogen bonds between the peptide and water and 10 between the peptide and the enzyme. PW1 appears as more hydrophilic, while PW2 is intrinsically more stable with more favorable enzyme-peptide interaction and more hydrophobic. The gain in energy of the peptide–solvent, enzyme–solvent and complex–solvent contributions is compensated by a loss in the solvent–solvent contribution. Table V compares the molecular accessible surface areas of both models PW1 and PW2. With the exception of Tyr1, all other residues of the inhibitor are found more buried in model PW2, indicating a more extensive interaction with the active site of FXa for this model.

Further evidence in favor of model P2 or PW2 comes from the analysis of the quantitative structure–activity relationship (sequence requirement vs. inhibitory activity) of the whole group of this type of peptidic inhibitors.[13] As described in the introduction, the hydrophobic residue in the sequence position 2 (Ile or Leu) is essential for the inhibitory activity, and the peptides may even be truncated after Arg3. It must therefore be Leu4 that interacts specifically with the site S4, as observed in model P2/PW2. The less important Leu2 is exposed to the solvent contributing mainly to the formation of a compact fold and giving fewer interactions with FXa.

The short tripeptide, peptide B, is superimposed to peptide A in model PW2 for the backbone and the centroid of the first two residues in Figure 8. It can be seen that the first three residues in the two peptides fold in a very similar conformation and the hydrophobic cyclohexyl side chain, replacing Ile2 in peptide A, fits well in the S4 pocket.

We conclude, from all the above observations that the model P2/PW2 reflects the reality better than the model P1/PW1, since the 3D structure fits better the shape of the active site, energetic considerations are all in its favour and it is in agreement with the analysis of the quantitative structure–activity relationship. The model also accommodates both peptides A and B in the same way. In this model the peptide itself forms a sharp turn at the junction of the C- and N-terminal fragments at Arg3, which may be classified as C10 turn[34,35] with a hydrogen bond between the Tyr1 carbonyl and Ile4 NH. The turn confers the inhibitor a wedge-shaped form filling the cavity of the active site.

Mode of Action and Specificity

Based on the primary structure alone, the inhibitors might appear to be substrates of FXa. The biological tests as well as the long-term stability of their NMR spectra, however, reveal that they are not even very poor substrates of the FXa. The models, depicted in Figures 7 and 8, show that the inhibitors occupy the substrate pockets S1 and S4 of the active site of the enzyme but that their position indeed differs from that of a substrate. The peptide bond which would be hydrolyzed in a substrate (Arg3-Ile4 in peptide A and Arg3-$NH_2$4 in peptide B) is out of reach of the catalytic triad His57, Asp102, Ser195. The decisive feature of the inhibitory specificity is the hydrophobic interaction of the site S4 with the two residues at positions 1 and 2. The aliphatic residue 2 is locked in site S4 and residue 1 closes the pocket like a lid, interacting with the enzyme residues Tyr99 and Phe174 as well as with the inhibitor residue 2. The enzymes for which the activity of the inhibitors has been found to be nil or very low are given in Table VII. They are all very similar in their sequence and consequently in their biochemical properties. Most have a well-developed S4 hydrophobic site with the possible exception of trypsin, where the S4 pocket is much looser, and kallikrein, with a large insertion at this position. The S1 pocket is nearly identical, especially in the enzymes preferring arginine to lysine.

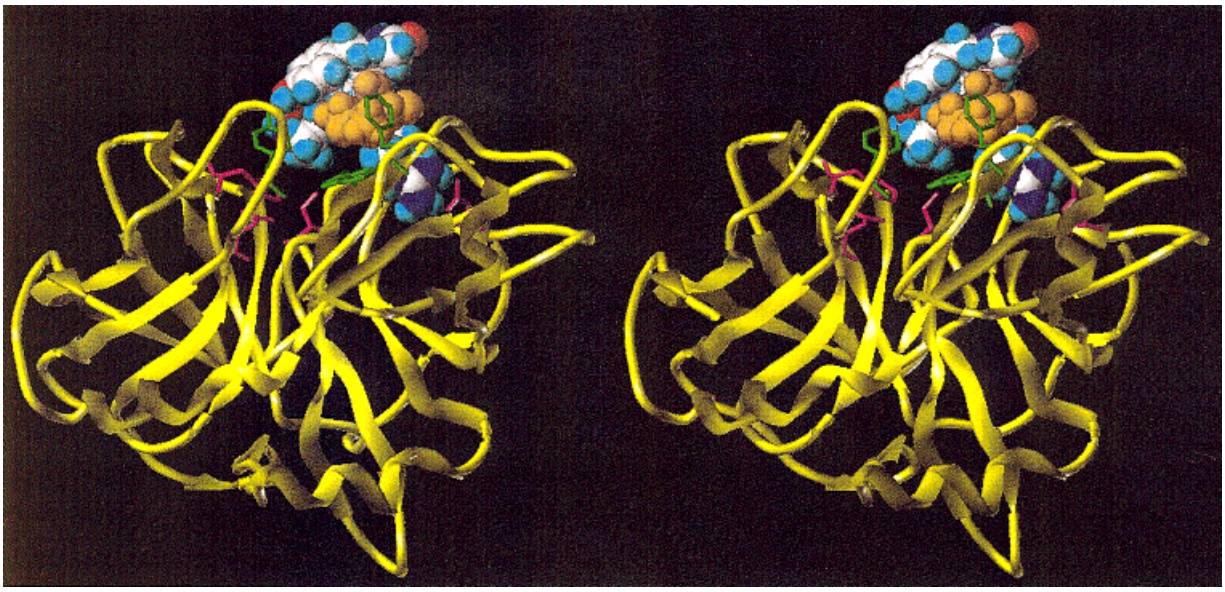

Fig. 7. Stereoview of P2 model of peptide A obtained after systematic search, SA and EM inside the active site of factor Xa. The enzyme is shown in yellow, the catalytic triad (His57, Asp102, and Ser195) and the residue anchoring the inhibitor Arg3 on the bottom of the pocket S1 (Asp189) are highlighted in magenta. The hydrophobic pocket S4 (Tyr99, Phe174, Trp215) is green. The peptide is in space-filling representation, with atom type coloring, except the Ile2 residue highlighted in orange.

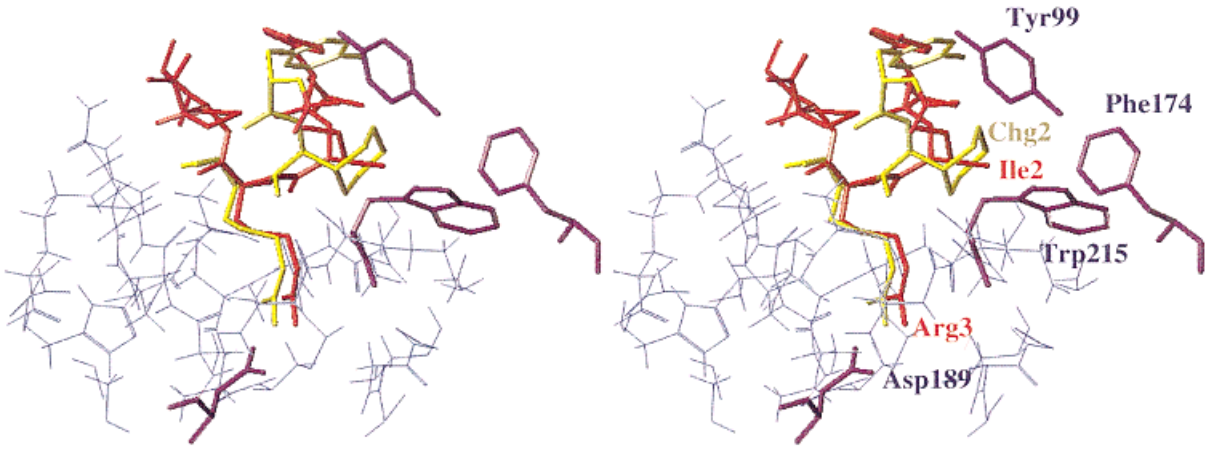

Fig. 8. Stereoview of model PW2 (final model), peptide A (*yellow*) overlaid with peptide B (*red*) inside the active site of FXa. Enzyme atoms within 10 Å radius from C^{α} of peptide A Arg3 are displayed (violet), S4 residues (Tyr99, Phe174, and Trp215) and the S1 residue Asp189) are highlighted (green).

The models indicate that the specificity resides in the S4 pocket. This pocket was suggested as the specificity site for the so-called "Daiichi compounds," a very different group of low molecular weight organic molecules acting as specific inhibitors of FXa.[6,8,9] A recent X-ray structure, published during the edition of the present manuscript,[36] confirms the importance of the S4 pocket for the "Daiichi compounds." The amino acid sequences in the vicinity of the S4 site are aligned with the sequence of FXa in Table VII. The only strictly conserved residue is Trp215 at the bottom of the pocket. The variation of aromatic, polar or apolar aliphatic amino acids at position 99 has been noted.[3,5,7] It should be also stressed that the greatest variability is confined to position 174. It may be the subtle combination of the properties of the residues at position 99 and 174 which is responsible for the remarkable specificity of the inhibitors studied here.

CONCLUSIONS

Both inhibitors A and B assume a compact 3D structure well adapted to the shape of the active site of FXa. They are wedged in the active site of the enzyme in a way differing from the binding of a substrate, hindering the access of the catalytic resi-

dues to the cleavable peptide bond. Arg3 is inserted in the pocket S1. The aliphatic hydrophobic side chain in the sequence position 2 is enveloped by the first residue in the inhibitor sequence and by residues Tyr99, Phe174, and Trp215 of the site S4. The selectivity is exerted above all by the site S4. Considering the broad similarity of the blood clotting proteinases, the fine differences in the S4 site seem to be a very attractive target in the design of further selective inhibitors not only for FXa, but also for other enzymes of blood coagulation cascade.

ACKNOWLEDGMENT

We thank Herman Schreuder for very useful discussions and the access to unpublished X-ray structures of inhibited thrombin and trypsin and Gert Vriend for help in the calculations of the molecular accessible surface areas.

REFERENCES

1. Lesk, A.M., Fordham, W.D. Conservation and variability in the structures of serine proteinases of the chymotrypsin family. J. Mol. Biol. 258:501–537, 1996.
2. Davie, E.W., Fujikawa, K., Kisiel, W. The coagulation cascade: Initiation, maintenance and regulation. Biochemistry 30:10363–10370, 1991.
3. Banner, D.W., D'Arcy, A., Chene, C., Winkler, F.K., Guha, A., Konigsberg, W.H., Nemerson, Y., Kirchhofer, D. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. Nature 380:41–46, 1996.
4. Brandstetter, H., Bauer, M., Huber, R., Lollar, P., Bode, W. X-ray structure of clotting factor IXa: Active site and module structure related to Xase activity and hemophilia B. Proc. Natl. Acad. Sci. U.S.A. 92:9796–9800, 1995.
5. Padmanabhan, K., Padmanabhan, K.P., Tulinsky, A., Park, C.H., Bode, W., Huber, R., Blankenship, D.T., Cardin, A.D., Kisiel, W. Structure of human des(1-45) factor Xa at 2.2 Å resolution. J. Mol. Biol. 232:947–966, 1993.
6. Stubbs, M.T., Huber, R., Bode, W. Crystal structures of factor Xa specific inhibitors in complex with trypsin: Structural grounds for inhibition of factor Xa and selectivity against thrombin. FEBS Lett. 375:103–107, 1995.
7. Rezaie, A.R. Role of residue 99 at the S2 subsite of factor Xa and activated protein C in enzyme specificity. J. Biol. Chem. 271:23807–23814, 1996.
8. Katakura, S., Nagahara, T., Hara, T., Iwamoto, M. A novel factor Xa inhibitor: Structure-activity relationships and selectivity between factor Xa and thrombin. Biochem. Biophys. Res. Comm. 197:965–972, 1993.
9. Lin, Z., Johnson, M.E. Proposed cation-π mediated binding by factor Xa: A novel enzymatic mechanism for molecular recognition. FEBS Lett. 370:1–5, 1995.
10. Herzberg, M. Biochemistry of factor X. Blood Rev. 8:56–62, 1994.
11. Waxman, L., Smith, D.E., Arcuri, K.E., Vlasuk, G.P. Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. Science 248:593–596, 1990.
12. Vlasuk, G.P., Ramjit, D., Fujita, T., Dunwiddie, C.T., Nutt, E.M., Smith, D.E., Shebuski, R.J. Comparison of the in vivo anticoagulant properties of standard heparin and the highly selective factor Xa inhibitors antistasin and tick anticoagulant peptide (TAP) in a rabbit model of venous thrombosis. Thromb. Haemost. 65:257–262, 1991.
13. Ostrem, J.A., Al-Obeidi, F.A., Safar, P., Safarova, A., Stringer, S.K., Patek, M., Cross, M.T., Spoonamore, J., LoCascio, J.C., Kasireddy, P., Thorpe, D., Sepetov, N., Lebl, M., Wildgoose, P., Strop, P. Discovery of a novel, potent, and specific family of factor Xa inhibitors via combinatorial chemistry. Submitted for publication.
14. Lam, K.S., Salmon, S.E., Hersch, E.M., Hruby, V.J., Kazmierski, W.M., Knapp, R.J. A new type of synthetic peptide library for identifying ligand-binding activity. Nature 354: 82–84, 1991.
15. Clore, G., Gronenborn, A.M. Theory and applications of the transferred nuclear Overhauser effect to the study of the conformations of small ligands bound to proteins. J. Magn. Reson. 48:402, 1982.
16. Ni, F. Recent developments in transferred NOE methods. Prog NMR Spectrosc 26:517–606, 1994.
17. Marion, D., Ikura, M., Tschudin, R., Bax, A. Rapid recording of 2D NMR spectra without phase cycling: Application to the study of hydrogen exchange in proteins. J. Magn. Reson. 85:393–399, 1989.
18. Briand, J., Ernst, R.R. Computer optimized homonuclear TOCSY experiments with suppression of cross relaxation. Chem. Phys. Lett. 185:276–285, 1991.
19. Piotto, M., Saudek, V., Sklenar, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J. Biomol. NMR 2:661–665, 1992.
20. Sklenar, V., Piotto, M., Leppik, R., Saudek, V. Gradient tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity. J. Magn. Reson. Ser. A 102:241–245, 1993.
21. Sklenar, V. Suppression of radiation damping in multidimensional NMR experiments using magnetic field gradients. J. Magn. Reson. Ser. A 114:132–135, 1995.
22. Desvaux, H., Berthault, P., Birlirakis, N., Goldman, M., Piotto, M. Improved versions of off-resonance ROESY. J. Magn. Reson. Ser. A 113:47–52, 1995.
23. Wüthrich, K. "NMR of Proteins and Nucleic Acids." New York: Wiley, 1986.
24. Clark, M., Cramer III, R.D., Van Opdenbosch, N. Validation of the general purpose Tripos 5.2 force field. J. Comp. Chem. 10:982–1012, 1989.
25. Weiner, S.J., Kollman, P.A., Nguyen, D.T., Case, D.A. An all atom force field for simulations of proteins and nucleic acids. J. Comp. Chem. 7:230, 1986.
26. Güntert, P., Braun, W., Wüthrich, K. Efficient computation of 3D protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. J. Mol. Biol. 217:517–530, 1991.
27. Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. Direction set (Powell's) methods in multidimensions. In "Numerical Recipes in C: The Art of Scientific Computing." Cambridge, U.K.: Cambridge University Press, 1988: 294–300.
28. Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. Downhill Simplex method in multidimensions. In "Numerical Recipes in C: The Art of Scientific Computing." Cambridge, U.K.: Cambridge University Press, 1988:290–293.
29. Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. J. Comp. Phys. 23:327, 1977.
30. Blanco, M. Molecular Silverware. I. General solutions to excluded volume constrained problems. J. Comp. Chem. 12:237, 1991.
31. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81:3684–3690, 1984.
32. Vriend, G. WHAT IF: A molecular modeling and drug design program. J. Mol. Graphics 8:52–55, 1990.
33. Arepalli, S.R., Glaudemans, C.P.J., Daves, G.D., Kovac, P., Bax, A. Identification of protein-mediated indirect NOE effects in a disaccharide-Fab complex by transferred ROESY. J. Magn. Reson. 106:195–198, 1995.
34. Benedetti, E. Structure and conformation of peptides as determined by X-ray crystallography. Chem. Biochem. Amino Acids Peptides Proteins 6:105–184, 1981.
35. Rose, G.D., Gierasch, L.M., Smith, J.A. Turns in peptides and proteins. Adv. Protein Chem. 37:1–109, 1985.
36. Brandstetter, H., Kühne, A., Bode, W., Huber, R., von der Saal, W., Wirthensohn, K., Engh, R.A. X-ray structure of active site-inhibited clotting factor Xa. J. Biol. Chem. 271:29988–29992, 1996.