Bovine Thrombin Complexed with an Uncleavable Analog of Residues 7−19 of Fibrinogen Aα: Geometry of the Catalytic Triad and Interactions of the P1′, P2′, and P3′ Substrate Residues[†,‡]

Philip D. Martin,[§] Michael G. Malkowski,[§] John DiMaio,[∥] Yasuo Konishi,[∥] Feng Ni,[∥] and Brian F. P. Edwards*[,§]

Department of Biochemistry, Wayne State University, Detroit, Michigan 48201, and Biotechnology Research Institute, National Research Council of Canada, 6100 Avenue Royalmount, Montréal, Québec, Canada

Received March 18, 1996; Revised Manuscript Received July 3, 1996[⊗]

ABSTRACT: The crystal structure of the noncovalent complex of bovine thrombin and a fibrinogen-Aα tridecapeptide substrate analog, G17ψ, in which the scissile bond amide nitrogen of Gly-17f has been replaced by a methylene carbon, has been determined at 2.3 Å resolution with an *R* factor of 17.1%. The geometry of the active site indicates that the crystal structure is a close model of the true Michaelis complex. The three independently determined thrombin/G17ψ complexes in the crystal asymmetric unit reveal novel interactions for the P2′ and P3′ residues—Pro-18f and Arg-19f, respectively—on the carboxyl-terminal side of the scissile bond and confirm previously observed interactions of the P1 (Arg-16f) through P10 (Asp-7f) positions on the amino-terminal side. The thrombin S2′ binding site for Pro-18f, as observed in all three complexes, differs from that predicted by modeling studies and is notable for including two carbonyl oxygens of the thrombin main chain. Arg-19f occupies two binding sites on thrombin, S3′A and S3′B, which have dramatically different placements for the arginyl side chain and carboxyl terminus.

The multifunctional role of thrombin (EC 3.4.21.5), as both participant and regulator, in the blood clotting cascade is well established (Berliner, 1992). Thrombin is a serine protease composed of a 49-residue A-chain and a 259-residue B-chain linked by a disulfide bond (MacGillivray & Davie, 1984; Magnusson et al., 1975) that cleaves fibrinogen at the Arg-16f/Gly-17f [1] bond of the Aα-chain and the Arg-14f/Gly-15f bond of the Bβ-chain. Crystal structures have been determined for bovine and human thrombin complexed with an wide assortment of peptide-based ligands. The structures of bovine thrombin complexed with residues 7f-16f of fibrinopeptide-A (Martin et al., 1992), hirudin (Vitali et al., 1992), and a hirudin C-terminal peptide (Vitali et al., 1996) have been determined. Similarly, structures of human thrombin covalently labeled with peptide chloromethyl ketones (Bode et al., 1989; Stubbs & Bode, 1992) or bound noncovalently to hirudin (Grütter et al., 1990; Rydel et al., 1990) or to hybrids of active site and C-terminal hirudin peptides (Qiu et al., 1992; Skrzypczak-Jankun et al., 1991; Zdanov et al., 1993) have also been reported. The structures of other serine proteases complexed with various inhibitors are known [reviewed in Polgar (1989)], and their relevance to the mechanism has been discussed (Schechter & Berger, 1967; Warshel et al., 1989).

A primary function of thrombin is to convert fibrinogen, which is a heterodimer of three chains ([Aα, Bβ, γ]$_2$) to insoluble fibrin ([α, β, γ]$_2$) by cleaving fibrinopeptides A and B, respectively, from the intact Aα- and Bβ-chains (Blomback et al., 1967; Hogg & Blomback, 1978). In addition, thrombin is both a positive and negative regulator of coagulation. Thrombin activates factors V and VIII, which are needed in prior steps of the blood clotting cascade and factor XIII, which covalently cross-links fibrin. Thrombin also stimulates platelet aggregation by activation of a specific platelet receptor (Vu et al., 1991). Thrombin is a negative regulator when bound to thrombomodulin, a protein cofactor that makes protein C the preferred substrate of thrombin. After cleavage by thrombin, protein C inactivates factors Va and VIIIa (Esmon, 1989).

Peptide mapping (Blombäck, 1986; Scheraga, 1986) and mutational studies (Le Bonniec et al., 1991) have shown that the P1′−P4′ residues (Figure 1) of substrates interact with thrombin and affect the rate of reaction. We report here the refined crystal structure of bovine thrombin complexed noncovalently with a substrate analog, G17ψ[2] (Figure 1), a peptide comprising residues 7f−19f of the Aα-chain of fibrinogen with a noncleavable C(O)−CH_2 linkage replacing the Arg-16f/Gly-17f peptide bond. This psuedo-Michaelis complex is the first thrombin structure to have Pro-18f and Arg-19f, respectively, in the P2′ and P3′ positions, as in

[†] This work was supported in part by NIH Grant GM 33192 (B.E.), a fellowship (M.G.M.) from NIH Training Grant T32 HL07602, and a Grant (F.N.) from the Medical Research Council of Canada (MRC-MT-12566).

[‡] The coordinates of the G17ψ complexes and related structure factors have been deposited with the Brookhaven Protein Data Bank (file names 1UCY and R1UCYSF, respectively).

* Corresponding author. Telephone: (313) 577-1506. Fax: (313) 577-2765. E-mail: bedward@cms.cc.wayne.edu.

[§] Wayne State University.

[∥] National Research Council of Canada.

[⊗] Abstract published in *Advance ACS Abstracts,* September 15, 1996.

[1] Sequence numbers for fibrinopeptide residues and hirudin residues are distinguished from those for thrombin residues by the suffix "f" and "h", respectively. Thrombin residues are numbered by homology with chymotrypsin (Bode et al., 1989).

[2] Abbreviations: E/SA, enzyme/substrate analog complex; E/P, enzyme/product complex; G17ψ , residues 7f−19f of the human fibrinogen Aα-chain with Gly-17f replaced with pseudoglycine [-$C(H)_2$-$C(H)_2$-C(O)-]; FpA7, residues 7f−16f of the human fibrinogen Aα-chain; PPACK, D-phenylalanyl-L-prolyl-L-arginyl chloromethylketone; rms, root mean square; RSfit, real space correlation coefficient; Agl, a combined arginine/pseudoglycine residue.

S0006-2960(96)00656-3 CCC: $12.00

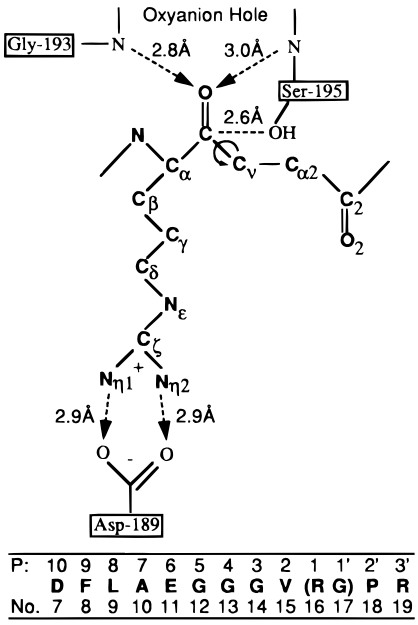

FIGURE 1: The G17ψ peptide analog. (Top) The covalent structure of the "Agl" residue and its major interactions with thrombin are shown in a diagram with standard labels for the Agl atoms. The Greek symbols are transliterated as α, A; β, B; γ, G; δ, D; ϵ, E; ζ, Z; η, H; and ν, N in the coordinate list (Protein Data Bank format) used for later figures. The ketone oxygen is bound in the oxyanion hole, close to the Oγ atom of Ser-195, and the arginyl side chain is pinned at the bottom of the S1 specificity pocket. The circular arrow indicates the new rotatable single bond in the Agl analog that replaces the rigid peptide bond in the native substrate. (Bottom) The amino acid sequence of G17ψ in the one-letter code is shown with the substrate numbers ("P"), as defined in Schechter and Berger 1967), above and the residue numbers ("No.") in the Aα-fibrinogen sequence below. The Agl residue replaces Arg-16f and Gly-17f (in parentheses). The amino terminus of Asp-7f is acetylated.

normal fibrinogen. The three, independently determined thrombin/G17ψ complexes (labeled I, II, and III) in the crystal asymmetric unit confirm the interactions of the P1 (Arg-16f) through P10 (Asp-7f) positions on the amino-terminal side (Martin et al., 1992) and reveal novel interactions for the P′ residues on the carboxyl-terminal side of the scissile bond.

MATERIALS AND METHODS

Data Collection. Bovine thrombin and the G17ψ peptide were prepared and cocrystallized as previously described (DiMaio et al., 1991, 1992; Martin et al., 1992). The crystals are isomorphous with the enzyme/product complex crystals (Martin et al., 1992) and belong to space group $P2_1$ with unit cell dimensions $a = 82.5$ Å, $b = 88.8$ Å, $c = 98.6$ Å, and $\beta = 106.2°$. There are three separate complexes of bovine thrombin and the fibrinopeptide in the asymmetric unit. Data were collected from one crystal with a Siemens X1000 area detector, three axis goniometer, Rigaku RU200H Cu X-ray source (40 kV, 70 mA), and a Supper graphite monochromator. Three "sweeps" of 110° in ω were collected (440 frames of data, 0.25°/frame, at least 10^6 counts/frame); the ϕ angle was incremented by 70° between sweeps. The data were merged and scaled with the XENGEN package (Howard et al., 1987) yielding 117 995 reflections with $I/\sigma > 1.0$ of which 45 875 were unique between 7.0 and 2.2 Å and used in the structure analysis. These data, which included 81% of the possible reflections in the 7.0−2.4 Å shell, 35% in the 2.4−2.3 Å shell, and 15% in the 2.3−2.2 Å shell, had an unweighted absolute R value based on intensity (R_{sym}) of 0.072.

Refinement of the Structure. Refinement was initiated by using the rigid body option in X-PLOR (Brünger, 1988), 7.0−2.5 Å data, and each of the independent thrombin and G17ψ molecules in the asymmetric unit as a rigid body. Following this step, 17 cycles of restrained, least-squares refinement with PROFFT (Agarwal, 1978; Finzel, 1987; Hendrickson & Konnert, 1980) dropped the R factor from 0.29 to 0.24 using $2\sigma_F$ data between 7.0 and 2.2 Å. A $2F_o - F_c$ electron density "omit" map showed the P1′−P2′ residues in the active sites of all three complexes and P3′ in complexes I and III. Water molecules were then added at positions that were within stereochemically reasonable distances from hydrogen bonding donors or acceptors and had electron density in both $F_o - F_c$ and $2F_o - F_c$ maps. A total of 87 cycles with PROFFT dropped the R factor to 0.177.

Up to this point, a normal Arg-16f/Gly-17f peptide bond was used in the refinement, although it is actually a ketone moiety that is not required to be planar along the main chain. Therefore, we switched to positional refinement in X-PLOR (Brünger, 1988; Engh & Huber, 1991) and defined the "Agl" residue (Figure 1) as a arginine/glycine hybrid which kept the planarity restraints for the ketone [Cα−C(O)−Cν] but removed the "planar amide nitrogen" restraint on the Cν atom [C(O)−Cν−Cα2]. Electron density maps calculated at this time also revealed an alternate position for Arg-19f, which was included in further refinements using X-PLOR.

The average temperature factors for Arg-16f in complexes I, II, and III, which were 30.8, 33.8, and 12 Å^2, respectively, differed unreasonably, given that Arg-16f had similar binding interactions in all three complexes. As in the E/P crystals (Martin et al., 1992), this discrepancy was rectified by calculating fractional occupancies for the G17ψ analogs in complexes I and II. Refinement with X-PLOR using alternating cycles of temperature factor and occupancy refinement gave final occupancies of 0.70 for fibrinopeptide I and II and 1.0 for fibrinopeptide III. After these calculations, the average temperature factors for the Arg-16f side chain in the S1 specificity pocket were 13.1, 13.6, and 12.4 Å^2 for fibrinopeptides I, II, and III, respectively. The mean isotropic temperature factor for the three complexes is 33.9 Å^2. The final structure had 8184 atoms, including 733 water oxygens.

Structural Analysis. Real space correlation coefficients (RSfit) for residues in the final $2F_o - F_c$ electron map were calculated with the program O (Jones et al., 1991). The robustness of the the refined structure was further evaluated by subjecting the final structure, minus solvent atoms, to three independent simulated annealing runs in X-PLOR using three different seed numbers for the random assignment of starting velocities at 4000 K (Brünger, 1988). The WA value in these calculations was equal to the checkstage weight.

The coordinates for the crystal structures of human thrombin complexed with hirutonin-2 (entry 1IHS), hirutonin-6 (entry 1IHT), and hirulog-3 (entry 1ABJ) were obtained from the Protein Data Bank (Bernstein et al., 1977). Hydrogen bonds were calculated using the program QUANTA (Polygen Corporation, Waltham, MA). The criteria for a hydrogen bond are that the angles C═O···H, O···H−N, and H−N−C be greater than 90° and the distance N···O not

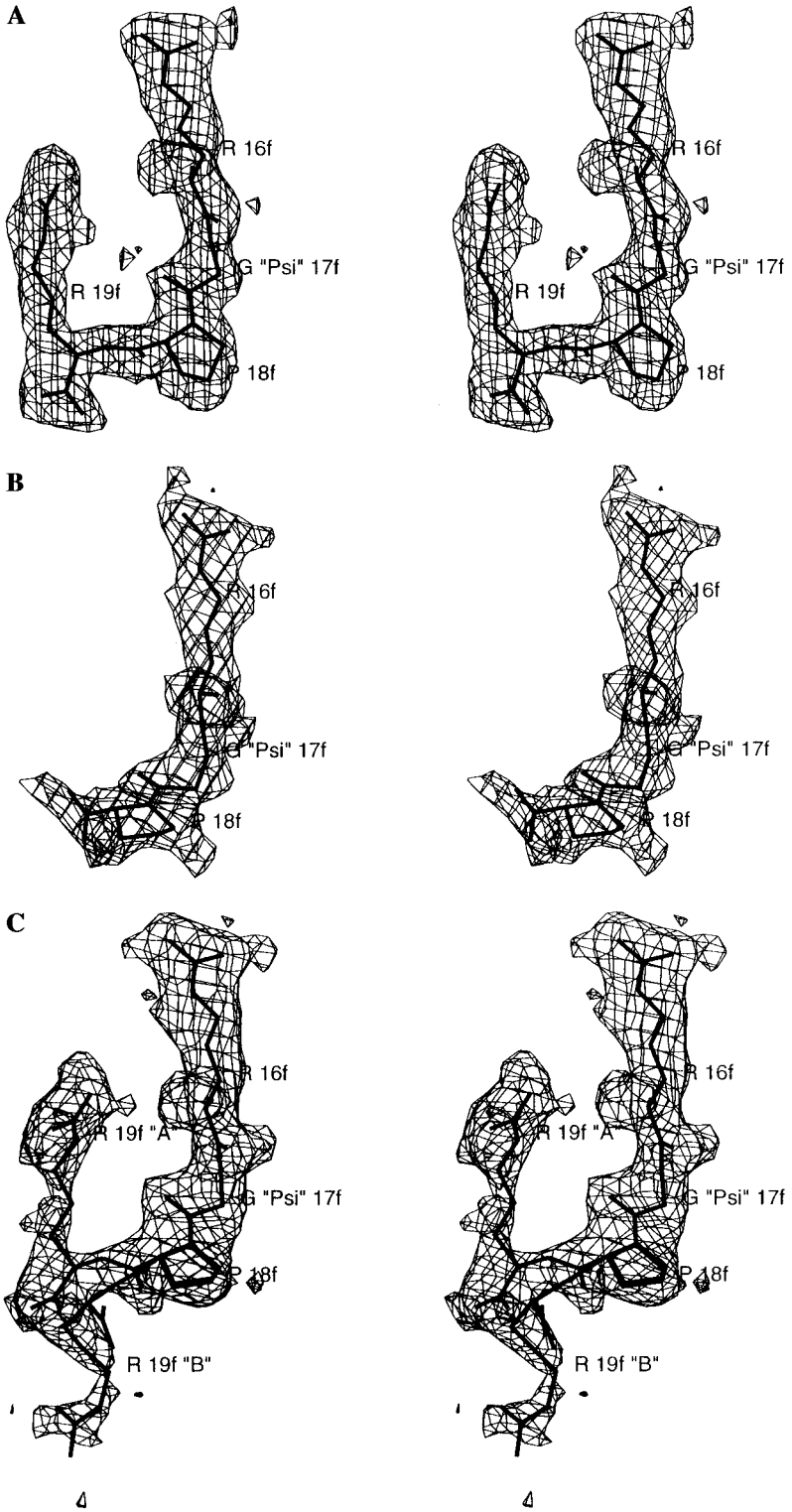

FIGURE 2: Stereo image of the $2F_o - F_c$ electron density for the P1−P3′ residues in G17ψ in (A) complex I, (B) complex II, and (C) complex III. The electron density within 2 Å of the atoms is contoured at 0.8σ for plots A and B and at 0.6σ for plot C where the density for Arg-19f is distributed between the S3′A and S3′B sites.

exceed 3.3 Å for a "regular" hydrogen bond and 4.0 Å for a "long" one (Baker & Hubbard, 1984).

RESULTS

The final *R* factor from PROFFT was 0.167 with rms deviations of 0.016 Å on bond distances and 1.8° on peptide bond torsion angles (ω). The final *R* factor after temperature factor refinement in X-PLOR was 0.171 with an estimated mean error in the coordinates of 0.25 Å from a Luzzati plot (Luzzati, 1952) and 0.29 Å from a σ_A plot (Read, 1986). RSfit is above 0.90 for the best defined residues of thrombin and as low as 0.30 for the worst defined residues, such as the C-terminus of the B-chain. In the Ramachandran plot calculated by PROCHECK (Laskowski et al., 1993), 97.9% of the residues fall within the allowed regions with a further 1.6% in the generously allowed regions and 0.5% in the disallowed regions. The non-glycine residues in the disallowed regions are from the ends of the A- and B-chains and

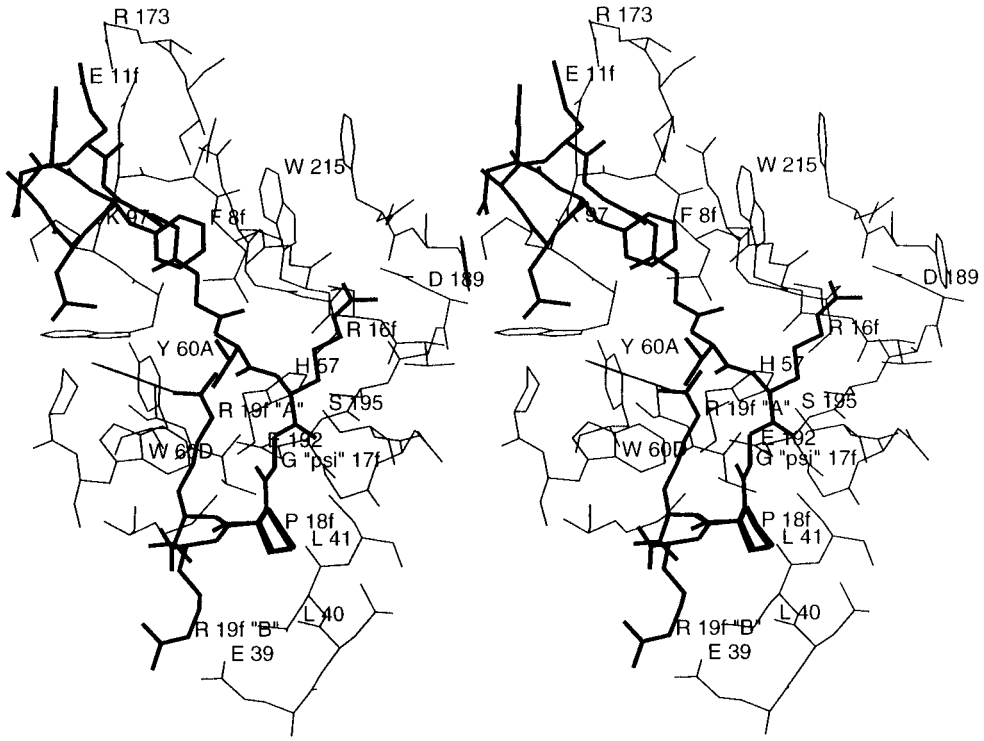

FIGURE 3: Interactions between G17ψ and thrombin. All thrombin residues (thin lines) within 4 Å of the G17ψ peptide (thick lines) in complex III are shown in stereo. The active site residues His-57, Asp-189, and Ser-195 fall within the 4 Å limit.

lie in poorly defined regions of the electron density map. Other measures of the stereochemistry, including the standard deviations for the ω angles (1.9°), ζ angles (1.8°), and pooled χ_1 angles (19.4°) are within the accepted range for other structures at a similar resolution (Laskowski et al., 1993).

The crystals have three crystallographically independent complexes of bovine thrombin/G17ψ in the asymmetric unit, labeled complex I, II, and III in this report. The same numbers are used for the respective thrombin molecules and bound G17ψ peptides. Complex I is adjacent to complex II, which in turn is related to complex III by a noncrystallographic two-fold axis. ϵ-Thrombin is generated when α-thrombin cleaves itself between residues Thr-149A and Ser-149B, in the so-called "autolysis loop", which is extremely mobile. Polyacrylamide electrophoresis (Laemmli, 1970) of the crystal used for data collection showed limited proteolysis of the thrombin molecule (about 40%) into ϵ-thrombin, similar to crystals of FpA7 bound to thrombin (Martin et al., 1992). As with the FpA7 crystals, there is density in molecule I through Thr-149A that stops and then starts again with residue Ser-149B and places the ends of the chains 15 Å apart, showing that molecule-I is ϵ-thrombin. In molecules II and III, on the other hand, the density is rather poor within the 149A and 149B segment, but well defined density on either side is easily spanned by a contiguous loop, indicating that molecules II and III are primarily α-thrombin.

The G17ψ peptide is well defined in complexes I, II, and III as indicated by average RSfit values of 0.78, 0.75, and 0.87, respectively, for the P9 to P1 residues. The electron density for the P1′ through P3′ residues of G17ψ is also well defined in all three complexes (Figure 2) with the exception of Arg-19f (P3′) in complex II which is disordered. We have adopted complex III as the reference structure because it is α-thrombin, because it retains full occupancy for the G17ψ peptide in the refinement calculations (see above), and, most importantly, because it has no crystal packing contacts near the S′ sites that might distort the binding of the peptide.

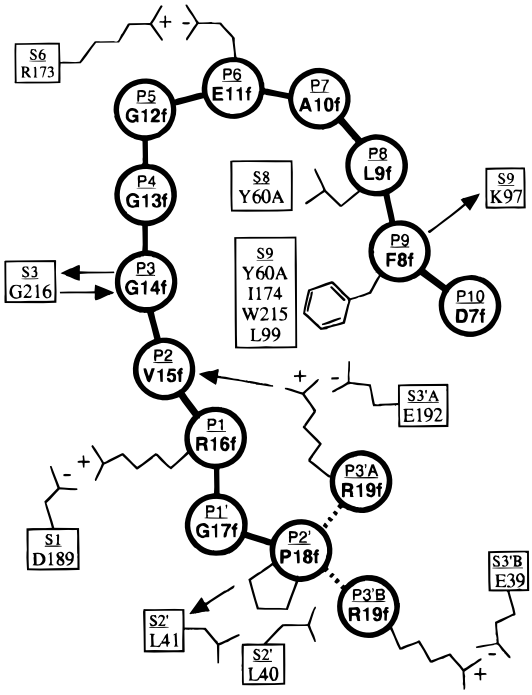

FIGURE 4: Cartoon showing the interactions between G17ψ and thrombin. Analog residues are enclosed in circles, thrombin residues in boxes. Sn and Sn′ are used for the binding sites on the enzyme for the Pn and Pn′ residues, respectively (Schechter & Berger, 1967). Hydrogen bonds are depicted as arrows from the hydrogen donor to the acceptor. The four thrombin residues in the S9 site are part of a hydrophobic cage around Phe-8f.

Interactions of P1−P9. The interactions in the enzyme/substrate analog complexes (E/SA) reported here (Figures 3, 4) are virtually identical with those discussed previously for the enzyme/product (E/P) complex with bound FpA7

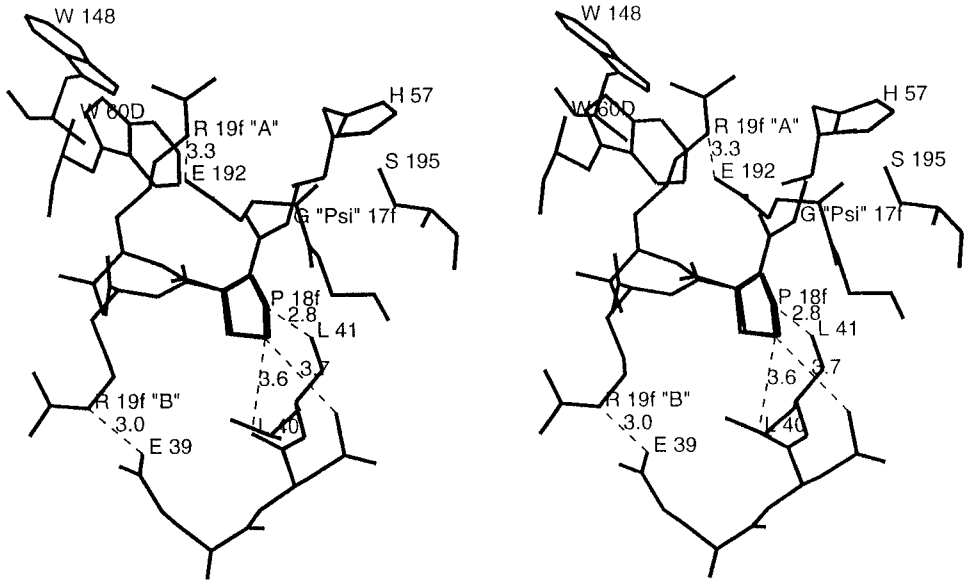

FIGURE 5: Observed S′ binding sites in complex III. Thrombin residues observed within 4 Å of "Gly17f", Pro-18f, and the two conformations of Arg-19f are shown. Contacts discussed in the text are identified with dashed lines and their distance (Å). The point of view is similar to that of Figure 3.

(Martin et al., 1992). The corresponding G17ψ and FpA7 peptides in the two crystal structures have rms deviations of 0.24 Å or less, based upon 32−37 main chain atoms (out of 40 total main chain atoms) judged equivalent by OVERLAP (Rossman & Argos, 1975). The identical calculation resulted in an rms deviation of 0.40 Å when the 93 side chain atoms alone were used. Any larger deviations that do occur are exclusively at the active site and are discussed below.

Interactions of P1′. There are no interactions, other than van der Waals, for the pseudoglycine residue in the P1′ site. The C_ν atom (Figure 1) which represents the amide nitrogen of an actual substrate that receives a proton from His-57 during proteolysis, is within 4.0 Å of His-57 Nϵ2 in all three complexes. Other natural thrombin substrates have valine, isoleucine, serine, or threonine at P1′ (Stubbs & Bode, 1993). These residues can be modeled comfortably into the S1′ site by changing only side chain torsion angles. Valine and isoleucine can establish hydrophobic contacts with Trp-60D, serine can easily form a hydrogen bond with Lys-60F, and threonine can potentially have its Cγ2 atom in hydrophobic contact with Trp-60D and its Oγ atom forming a hydrogen bond with Lys-60F.

Interactions of P2′. The C_γ and C_δ methylene groups of Pro-18f contact the carbonyl oxygen of Leu-40 (3.6 Å) and Leu-41 (2.8 Å), respectively (Figure 5), in all three complexes. These close contacts are retained in the three independent cool stage refinements with standard deviations of less than 0.2 Å. The proline Cγ carbon also makes a hydrophobic contact with the $C_{\delta1}$ of Leu-40 (3.7 Å). The carbonyl oxygen of Leu-41 is also close to a conserved water molecule (3.3, 2.7, and 2.8 Å) which is hydrogen bonded to Lys-60F Nζ in all three complexes. The P2′ residues of other thrombin substrates (Stubbs & Bode, 1993) can be modeled into this S2′ site without changing the backbone atoms or grossly disturbing the environment. Valine can occupy the same space as Pro-18f. Histidine can form simultaneous hydrogen bonds with Leu-41 (Nδ1···O), and Glu-39 (Nϵ2···Oϵ2), when protonated. Phenylalanine can potentially form a hydrophobic contact with the side chain atoms of Leu-40 and the C_α of Gly 193. Although glutamic acid cannot interact favorably with the S2′ residues, which are either neutral, hydrophobic, or negatively charged, its side chain can project into solvent.

Table 1: Interactions of the P3′ Residue, Arg-19f

	complex I[a]	distance (Å)	complex III	distance (Å)
polar interactions				
Arg-19f Nϵ	Wat-59	2.9	Glu-192 Oϵ1	3.3
			Glu-39 Oϵ2[b]	3.0
Arg-19f Nη1	Glu-192 Oϵ2	3.1	Glu-39 Oϵ2[b]	2.8
Arg-19f Nη2	Val-15f O	3.2	Val-15f O	2.8
nonpolar interactions				
Arg-19f Cγ	Trp-60D Cη2	3.8	Trp-60D Cη2	3.4
			Trp-60D Cζ2	3.9

[a] Complex I related by noncrystallographic symmetry. [b] The alternate Arg-19f position.

Interactions of P3′. Complex III has two S3′ binding sites for Arg-19f, labeled S3′A and S3′B (Figures 3 and 5). In complex I, the S3′B site is blocked by a crystal packing interaction between Glu-39 and Asn-204B in a symmetry-related thrombin. In complex II both S3′ sites are near Asp-222 in an adjacent complex and the electron density for Arg-19f is too weak to trace. S3′A is the same in complexes I and III and is stabilized by multiple hydrogen bonds or salt links, by a bridging water molecule (complex I only), and by hydrophobic interactions (Table 1).

The Geometry of the Catalytic Triad. The electron density and geometry for the catalytic site of complex III are shown in Figures 6 and 7. The plane of the scissile "peptide" bond is almost normal to a line drawn between the Oγ of Ser-195 and the ketone carbon of G17ψ as required for maximum activity (Klebe, 1991). The two atoms are closer (2.6 Å) than the sum of their van der Waals radii due to the arrested attack of the nucleophile on the carbonyl carbon of the ketone (Figure 7).

DISCUSSION

Canonical conformations for the six substrate residues, P3−P3′, have been deduced from complexes of serine proteases with small protein inhibitors (Bode & Huber, 1991,

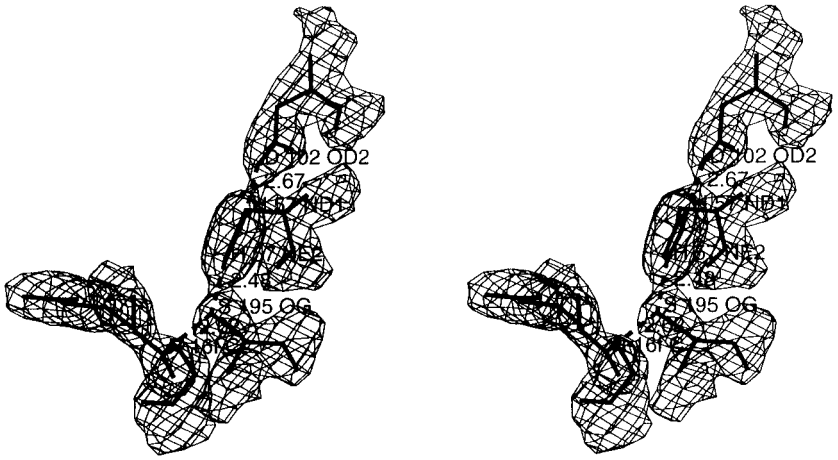

FIGURE 6: Active site in complex III. The $2F_o - F_c$ electron density for His-57, Asp-102, and Ser-195 in thrombin and Agl-16f in G17ψ is shown at the 2.0σ contour level. Distances for the Agl-C to Ser-195 Oγ, Ser-195 Oγ to His-57 Nδ1, and His-57 Nϵ2 to Asp-102 Oδ2 are 2.62, 2.48, and 2.67 Å, respectively.

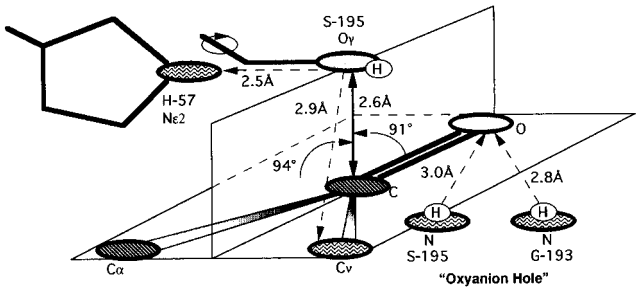

FIGURE 7: Reaction geometry of the catalytic site atoms in thrombin/G17ψ complex III. The diagram is modeled after Klebe (1991).

1992). The P2−P1′ residues of G17ψ have torsion angles mostly within the canonical range, but the P2′and P3′ residues do not (Table 2). The P′ sites have been modeled from the structure of thrombin complexed with hirulog-3, which contains D-Phe-Pro-homoarginine group linked via five glycines to a hirudin peptide in the exosite (Qiu et al., 1992). These results differ markedly from our observations (Figure 8) because the homoarginine residue offsets the P′ residues by one atom position toward the exosite and because the pentaglycine linker takes a "short-cut" to the exosite. The fibrinogen Aα-chain uses approximately 18 residues to loop out of the S3′ position, form an interchain disulfide bond, and then loop back into the exosite (Stubbs & Bode, 1993).

In contrast, the structure of hirutonin-2, which has PPACK linked to residues 48h−65h of hirudin by a pseudoglycine at P1′ (Zdanov et al., 1993), has essentially the same ϕ,ψ angles (−62°, 169°) at P1′ as G17ψ and traces a similar path through the S1′−S3′A subsites but then continues on to the exosite along the path of the side chain of Arg-19f in the S3′B site (Figure 9). Also, the plane of the peptide bond between the P2′ (Gln) and P3′ (Ser) residues in hirutonin-2 is flipped, placing the P2′ side chain in a different conformation.

The S′ Subsites. Although G17ψ has a methylene carbon, C_v, at P1′, it is still an excellent analog of the true substrate. The chemical differences introduced by the methylene carbon are (1) a partial charge of 0, as opposed to −0.35 for an amide nitrogen, (2) sp^3 hybridization with two hydrogen atoms attached instead of one for the planar sp^2 amide nitrogen, and (3) a C−C "scissile bond" that is longer (1.52 Å) and rotatable compared to a C−N peptide bond (1.34 Å). Despite these chemical differences, all three G17ψ complexes have values (166°, 172°, and 162°, respectively) for the torsion angle of the "scissile peptide bond" (C_α−C−C_ν−$C_{\alpha 2}$) that are close to a true substrate ($\omega = 180°$). For comparison, we modeled the geometry of the true substrate by inserting glycine in the P1′ position of G17ψ and doing 120 cycles of positional refinement using the X-PLOR force field with the X-ray terms. The corresponding Arg-16/Agl atoms were essentially identical (0.06 Å displacement between the Arg C atom and Agl C atom) while the corresponding Gly-17/Agl atoms showed displacements due to the shorter C−N peptide bond that decreased from 0.3 Å for the Gly N/Agl C_ν atoms to 0.09 Å for the Gly C/Agl C_2 atoms. Displacements for the corresponding Pro-18 and Arg-19 atoms were insignificant.

The S2′ site was not anticipated from previous work where the Pro-18f pyrrolidine ring was modeled in a depression formed by Glu-39, Leu-41, Cys-42−Cys-58, Trp-60D, and Lys-60F, with the Nζ of Lys-60F forming a hydrogen bond to the Pro-18f O (Stubbs & Bode, 1992). In the analog complexes, Glu-39, Cys-42, Cys-58, and Lys-60F are all greater than 4.0 Å from Pro-18F and not part of S2′, whereas newcomers Gly-193 and Leu-40 are part of S2′ (Figure 5). Electron density for the experimentally observed P2′ site is present in all three complexes (Figure 2). Although proline is classified as a hydrophobic residue, having Pro-18f locked firmly into an environment that includes the hydrophilic carbonyl oxygens of Leu-40 and Leu-41 is not unreasonable given that 40% of proline side chain contacts in proteins are with polar residues (Singh & Thornton, 1992).

The Cδ atom of Pro-18f is significantly closer to Leu-41 O (2.8 Å) than is the Cγ atom to Leu-40 O (3.6 Å). A weak, C−H···O=C< hydrogen bond between the Pro-18f Cδ atom and Leu-41 O is a possible, partial explanation of this difference. The Pro-18f Cδ atom, which is next to the electron-withdrawing imide nitrogen in the peptide bond, is approximately as acidic as the Cα atom of proline, which makes a C−H···O=C< hydrogen bond in the crystal structure of *N*-acetylleucylprolinamide (Puliti et al., 1992). Although weak, the C−H···O=C< hydrogen bond is documented in many small molecule structures (Desiraji, 1991; Jeffrey & Maluszynska, 1981; Taylor & Kennard, 1982) and also recently in protein crystal structures (Derewenda et al., 1994, 1995; Steiner & Saenger, 1993).

Table 2: Canonical Torsion Angles for P3 to P3′ [a]

		canonical[a]		complex I		complex II		complex III	
		ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ
P3 G14f		−120/−140	140/170	−135	(114)	(−144)	(124)	−138	(121)
P2 V15f		−60/−100	139/180	(−59)	151	−68	156	−67	157
P1 R16f		−95/−120	9/50	−109	27	−110	28	−106	29
P1′ G17f[b]		−60/−100	139/180	−73	178	(−57)	151	−65	175
P2′ P18f	"A"	−99/−140	70/120	(−70)	(143)	(−44)	101	(−76)	(132)
	"B"							(−83)	(−172)
P3′ R19f	"A"	−99/−140	70/120	−(98)	77[c]			(−95)	(175)[c]
	"B"							(−35)	95[c]

[a] The canonical values are "consensus" ranges deduced from complexes between proteases and small protein inhibitors (Bode & Huber, 1992). Torsion angles in the three thrombin/G17ψ complexes that fall outside the canonical ranges are enclosed in parentheses. [b] These angles were determined for a glycine refined in the P′1 position. [c] Measured from the terminal oxygen (OT1 or OT2) that gave an angle closest to the canonical value.

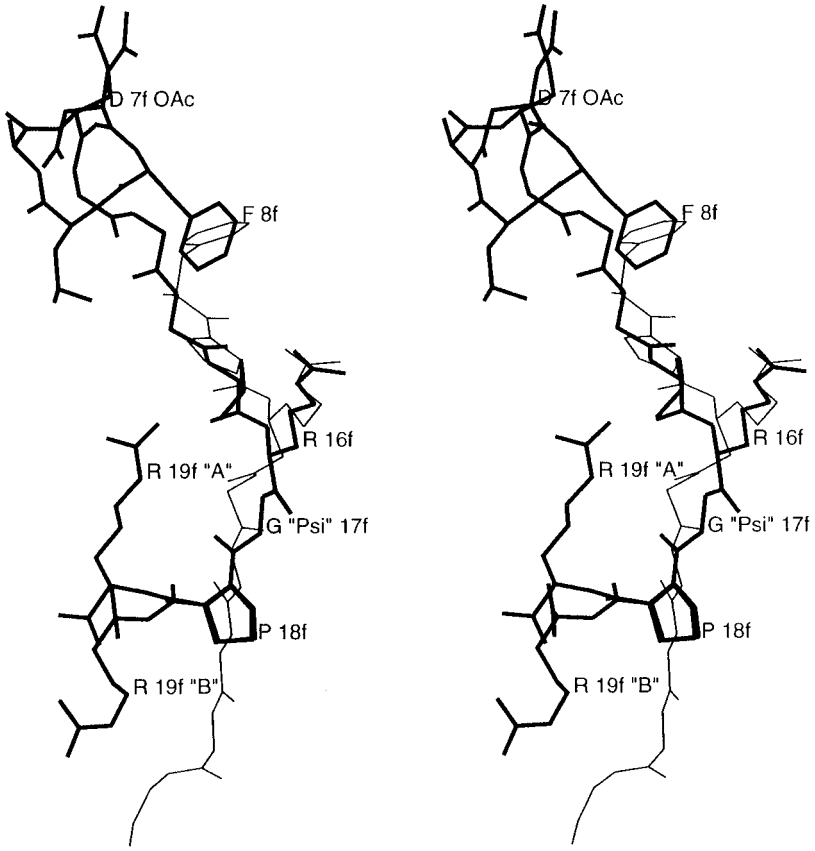

FIGURE 8: Comparison of G17ψ with hirulog-3. G17ψ (thick lines) in complex III and hirulog-3 (thin lines) were superimposed using the thrombin Cα atoms. The point of view is similar to that of Figure 3.

The observation of two S3′ sites in bovine thrombin/G17ψ complex III is the first direct, structural explanation for how thrombin can accommodate such radically different substrates as fibrinogen, which has arginine at P3′ and protein C, which has aspartate at P3′. The S3′ site that is common to complex I and complex III ideally suits fibrinogen and is named S3′A in these discussions. The site has favorable interactions for Arg-19f, which makes a hydrogen bond with its own backbone and with Glu-192. Moreover, a modeled P4′ residue goes out into solution, which is necessary to accommodate the 18-residue segment in fibrinogen that links the P′ residues to the residues bound in the exosite.

The enhanced reactivity of a E39K mutant thrombin toward protein C peptides identify Glu-39 as a major S3′ constituent for protein C proteolysis (Le Bonniec et al., 1991). The second site for Arg-19f, labeled S3′B, has the side chain pointing toward Glu-39 (Figures 5), as expected for protein C. Although charge repulsion from Glu-39 reduces binding of protein-C to S3′B in wild-type thrombin, the repulsion is alleviated in the complex with thrombomodulin (Le Bonniec et al., 1991).

The thrombin receptor has two bulky hydrophobic groups (Phe-Leu) at the P2′-P3′ position which would not be expected to interact favorably with thrombin. Indeed, in crystals containing the thrombin receptor peptide complexed to thrombin, residues P1−P3 in the active site and P10′−P14′ in the exosite are well defined but P1′−P9′ are not (Stubbs & Bode, 1993). Apparently, the interaction of this novel thrombin substrate with the exosite compensates adequately for the weak S1′−S4′ interactions.

Geometry of the Scissile Bond. The geometry of the catalytic residues in the G17ψ complex agrees with theoretical calculations for the attack of a nucleophile on a carbonyl carbon (Bürgi & Dunitz, 1983; Bürgi et al., 1973; Klebe,

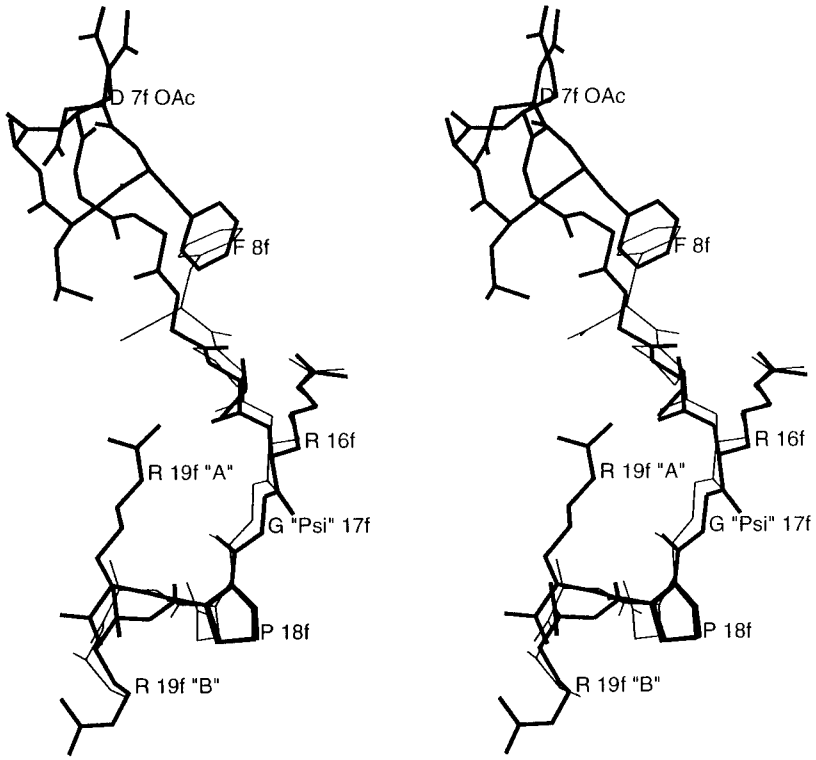

FIGURE 9: Comparison of G17ψ with hirutonin-2. G17ψ (thick lines) in complex III and hirutonin-2 (thin lines) were superimposed using the thrombin Cα atoms. The point of view is similar to that of Figure 3.

Table 3: Active Site Interactions[a]

	R16f/S195	H57/S195		H57/D102	
complex[a]	Oγ−C (Å)	Oγ−Nε2 (Å)	Cβ−Oγ−Nε2[b]	Nδ1−Oδ2 (Å)	Cγ−Oδ2−Nδ1
G17ψ I	3.0	2.4	131.7°	2.9	110.8°
G17ψ II	2.7	2.7	101.7°	2.6	108.7°
G17ψ III	2.6	2.5	112.9°	2.7	103.6°
FpA7 I	2.7	2.8	98.1°	2.7	127.1°
FpA7 II	2.4	2.8	92.0°	2.7	118.8°
FpA7 III	3.6[c]	2.9[c]	97.3°	2.5	130.0°

[a] The G17ψ values are for "Agl" at positions 16f and 17f. The FpA7 values are from (Martin et al., 1992). [b] The optimum value for maximum strength is 111° (Marquart et al., 1983). [c] Ser-195 is in an alternate conformation and forms a hydrogen bond with a backbone carbonyl oxygen.

Table 4: Reaction Geometry at the Active Site of the Thrombin/G17ψ Complexes[a]

complex	Oγ−C−O	Oγ−C−Cα	Oγ−C−Cν[b]	Oγ−C
G17ψ I	89.4° (91.7°)	91.9° (88.5°)	86.6° (91.7°)	3.1 Å (3.2 Å)
G17ψ II	86.3° (91.0°)	88.7° (86.0°)	90.4° (86.9°)	2.7 Å (2.6 Å)
G17ψ III[c]	90.8° (87.8°)	93.7° (89.4°)	83.1° (88.4°)	2.6 Å (2.7 Å)
ovomucoid	94.5°	83.8°	98.7°	2.7 Å

[a] The angles and distances are measured from the Oγ of Ser-195 to the Agl atoms in G17ψ (Figure 7) in a refinement that included both the "A" and "B" positions for peptide III. The standard deviations are estimated from the three "coolstage" refinements to be ±2° for the angles and 0.1 Å for the Oγ−C distance. The values in parentheses are for Arg-Gly, rather than Agl, refined in positions 16f and 17f. [b] The values in parentheses are for the corresponding Arg−Gly angle, O_γ−C−N. [c] When only the "A" position was included in the refinement, the values for peptide III were 87.9°, 93.3°, 86.4°, and 2.6 Å, respectively.

1991). If one constructs a reaction plane along the line of the C═O atoms and normal to the peptide bond plane, then the Oγ of Ser-195 in complex III is in the reaction plane and directly over the carbonyl carbon (Figures 6 and 7).

The Oγ of Ser-195 is compressed between the ketone carbon of Agl (i.e., the carbonyl carbon of Arg-16f) and His-57 Nε2, with distances between the two that would indicate some interpenetration of van der Waals spheres. Although the positions of the active site residues in the G17ψ complexes are well defined (Figure 6), as indicated by side chain *B* values that are lower (<20 Å^2) than the overall average temperature factor of 33 Å^2, the resolution of the data is not high enough to release the constraints on the ketone planarity and determine by refinement if the ketone carbon is distorted toward the tetrahedral intermediate (Marquart et al., 1983; Read et al., 1983). However, it is clear that the presence of the substrate analog favors a hydrogen bond between Ser-195 Oγ and His-57 Nε2 (Table 3). This hydrogen bond has less favorable angles in the E/P complex (Table 3) and is absent in uncomplexed bovine thrombin (unpublished) and other serine proteases (Bachovchin, 1986; Bode et al., 1983).

ε-Thrombin. Complex I contains ε-thrombin, which is α-thrombin cleaved at residue 149A in the autolysis loop, while complexes II and III are α-thrombin. ε-Thrombin has the same K_m value but only one-half of the k_{cat} value of α-thrombin toward the Aα-chain of fibrinogen (Hofsteenge et al., 1988). Because ε-thrombin has the same activity as α-thrombin toward small substrates, Hofsteenge and coworkers (1988) have concluded that the nucleophilicity of Ser-195 is unchanged and the reduced activity of ε-thrombin toward fibrinogen is "due to a less optimal alignment of the scissile bond relative to the catalytic residues". However, the G17ψ crystal stucture (Table 4) does not support this explanation in that the alignment of the scissile bond in ε-thrombin (complex I) is as close to ideal as in α-thrombin (complexes II and III). There are no structural differences larger than the estimated error of the model that satisfactorily explain the kinetic differences between α- and ε-thrombin.

The YPPW loop has moved about 1 Å in complex I relative to complex III, but this distortion can also be explained by crystal packing contacts.

Inhibitors. Protease B (*Streptomyces griseus*) complexed with the third domain of turkey ovomucoid is a text book example of a serine protease/protein inhibitor complex in which the scissile bond is intact (Read et al., 1983). Read et al. (1983) attribute the lack of reaction to the rigidity of the ovomucoid "substrate". However, the substrate analog G17ψ is also held rigidly in the active site, as shown by a low average temperature factor (13 Å^2) for Arg-16f in complex III. Furthermore, the native equivalent of G17ψ, with glycine at position 17f, is cleaved in the crystal (Martin et al., 1992) despite the fact that it is presumably held just as rigidly as the G17ψ analog. A more comprehensive explanation derived from this work is that peptide inhibitors like turkey ovomucoid react very slowly because they are held rigidly at the active site *with nonideal geometry*. As evident in Table 4, the three angles in the ovomucoid complex that define the approach of Ser-195 Oγ to the carbonyl carbon are farther from the ideal value of 90° for a nucleophilic attack (Klebe, 1991) than are the corresponding angles in the G17ψ substrate analog, even though the latter has a rotatable C−Cν bond instead of a peptide bond.

ACKNOWLEDGMENT

We thank Dr. Jacqueline Vitali for drawing our attention to the literature on C−H hydrogen bonds.

REFERENCES

Agarwal, R. C. (1978) A new least-squares refinement technique based on the fast fourier transform algorithm., *Acta Crystallogr. A34*, 791−809.

Bachovchin, W. W. (1986) ^{15}N NMR spectroscopy of hydrogen-bonding interactions in the active site of serine proteases: evidence for a moving histidine mechanism, *Biochemistry 23*, 7751−7759.

Baker, E. N., & Hubbard, R. E. (1984) Hydrogen bonding in globular proteins, *Prog. Biophys. Mol. Biol. 44*, 97−179.

Berliner, J. (1992) *Thrombin: Structure and Function*, Plenum Press, New York.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, G. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) The protein data bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol. 112*, 535−542.

Blombäck, B. (1986) Specificity of thrombin and its action on fibrinogen, *Ann. N. Y. Acad. Sci. 485*, 120−123.

Blomback, B., Blomback, M., Hessel, B., & Iwanaga, S. (1967) Structure of N-terminal fragments of fibrinogen and specificity of thrombin, *Nature 215*, 1445−1448.

Bode, W., & Huber, R. (1991) Ligand binding: proteinase−protein inhibitor interactions, *Curr. Opin. Struct. Biol. 1*, 45−52.

Bode, W., & Huber, R. (1992) Natural protein proteinase inhibitors and their interaction with proteinases, *Eur. J. Biochem. 204*, 433−451.

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G., & Bartunik, H. (1983) Refined 2 Å X-ray crystal structure of porcine pancreatic kallikrein A, a specific trypsin-like serine protease, *J. Mol. Biol. 164*, 237−282.

Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. R., & Hofsteenge, J. (1989) 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. 8*, 3467−3475.

Brünger, A. T. (1988) Crystallographic refinement by simulated annealing: application to a 2.8 Å resolution structure of aspartate aminotransferase, *J. Mol. Biol. 203*, 803−816.

Bürgi, H. B., & Dunitz, J. D. (1983) From crystal statics to chemical dynamics, *Acc. Chem. Res. 16*, 153−161.

Bürgi, H. B., Dunitz, J. D., & Shefter, E. (1973) Geometrical reaction coordinates II. Nucleophilic addition to a carbonyl group, *J. Am. Chem. Soc. 95*, 5065−5067.

Derewenda, Z. S., Derewenda, U., & Kobos, P. M. (1994) (His)Cϵ−H···O═C< Hydrogen bond in the active site of serine hydrolases, *J. Mol. Biol. 241*, 83−93.

Derewenda, Z. S., Lee, L., & Derewenda, U. (1995) The occurrence of C−H···O hydrogen bonds in proteins, *J. Mol. Biol. 252*, 248−262.

Desiraji, G. R. (1991) The C−H···O hydrogen bond in crystals: What is it?, *Acc. Chem. Res. 24*, 290−296.

DiMaio, J., Ni, F. G. B., & Konishi, Y. (1991) A new class of potent thrombin inhibitors that incorporates a scissile psuedopeptide bond, *FEBS Lett. 282*, 47−52.

DiMaio, J., Gibbs, B., Lefebvre, J., Konishi, Y., Munn, D., Yue, S. Y., & Hornberger, W. (1992) Synthesis of a homologous series of ketomethylene arginyl psuedodipeptides and application to low molecular weight hirudin-like thrombin inhibitors, *J. Med. Chem. 35*, 3331−3341.

Engh, R. A., & Huber, R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement, *Acta Crystallogr. A47*, 392−400.

Esmon, C. T. (1989) The roles of protein C and thrombomodulin in the regulation of blood coagulation, *J. Biol. Chem. 264*, 4743−4746.

Finzel, B. C. (1987) Incorporation of fast Fourier transforms to speed restrained least-squares refinement of protein structures, *J. Appl. Crystallogr. 20*, 53−55.

Grütter, M. G., Priestle, J. P., Rahuel, J., Grossenbacher, H., Bode, W., Hofsteenge, J., & Stone, S. R. (1990) Crystal structure of the thrombin-hirudin complex: a novel mode of serine protease inhibition, *EMBO J. 9*, 2361−2365.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Eds.) pp 43−57, The Indian Academy of Sciences, Bangalore.

Hofsteenge, J., Braun, P. J., & Stone, S. R. (1988) Enzymatic properties of proteolytic derivatives of human α thrombin, *Biochemistry 27*, 2144−2151.

Hogg, D. H., & Blomback, B. (1978) The mechanism of the fibrinogen-thrombin reaction, *Thromb. Res. 12*, 953−964.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) The use of an imaging proportional counter in macromolecular crystallography, *J. Appl. Crystallogr. 20*, 383−387.

Jeffrey, G. A., & Maluszynska, H. (1981) A survey of hydrogen-bond geometries in the crystal structures of amino acids, *Int. J. Biol. Macromol. 4*, 173−185.

Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.

Klebe, G. (1991) in *The Application of Charge Density Research to Chemistry and Drug Design* (Jeffrey, G. A., & Piniella, J. F., Eds.) pp 287−318, Plenum Press, New York.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.

Le Bonniec, B. F., MacGillivray, R. T. A., & Esmon, C. T. (1991) Thrombin Glu-39 restricts the P′3 specificity to non acidic residues, *J. Biol. Chem. 266*, 13796−13803.

Luzzati, V. (1952) *Acta Crystallogr. 5*, 802−810.

MacGillivray, R. T. A., & Davie, E. W. (1984) Characterization of bovine prothrombin mRNA and its translation product, *Biochemistry 23*, 1626−1634.

Magnusson, S., Petersen, T. E., Sottrup-Jensen, L., & Claeys, H. (1975) in *Proteases and Biological Control* (Reich, E., Rifkin, D. B., & Shaw, E., Eds.) pp 123−149, Cold Spring Harbor Laboratory, New York.

Marquart, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983) The geometry of the reactive site and of the peptide groups

in trypsin trypsinogen and its complexes with inhibitors, *Acta Crystallogr. B39*, 480−490.

Martin, P. D., Robertson, W., Turk, D., Huber, R., Bode, W., & Edwards, B. F. P. (1992) The structure of residues 7−16 of the Aa-chain of human fibrinogen bound to bovine thrombin at 2.3 Å resolution, *J. Biol. Chem. 267*, 7911−7920.

Polgar, L. (1989) in *Mechanism of Protease Action* (Polgar, L., Eds.) pp 87−122, CRC Press, Boca Raton, FL.

Puliti, R., Mattia, C. A., & Lilley, T. H. (1992) Structure of *N*-acetyl-L-leucyl-L-prolinamide monohydrate, *Acta Crystallogr. C48*, 709−712.

Qiu, X., Padmanabhan, K. P., Carperos, V. E., Tulinsky, A., Kline, T., Maraganore, J. M., & Fenton, I. J. (1992) Structure of the hirulog-3 thrombin complex and nature of the S′ subsites of substrates and inhibitors, *Biochemistry 31*, 11689−11697.

Read, R. J. (1986) Improved fourier coefficients for maps using phases from partial structures with errors, *Acta Crystallogr. A42*, 140−149.

Read, R. J., Fujinaga, M., Sielecki, A. R., & James, M. N. G. (1983) Structure of the complex of *Streptomyces griseus* protease B and the third domain of the turkey ovomucoid inhibitor at 1.8 Å resolution, *Biochemistry 22*, 4430−4433.

Rossman, M. G., & Argos, P. (1975) A comparison of the heme binding pocket in globins and cytochrome b_5, *J. Biol. Chem. 250*, 7525−7532.

Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, R., & Roitsch, C. (1990) The structure of a complex of recombinant hirudin and human α-thrombin, *Science 249*, 277−283.

Schechter, J., & Berger, A. (1967) On the size of the active site in proteases I. Papain. *Biochem. Biophys. Res. Commun. 27*, 157−162.

Scheraga, H. A. (1986) Chemical basis of thrombin interactions with fibrinogen, *Ann. N. Y. Acad. Sci. 485*, 124−133.

Singh, J., & Thornton, J. M. (1992) *Atlas of protein side-chain interactions*, Vol. 2, pp 587−627, IRL Press, Oxford.

Skrzypczak-Jankun, E., Carperos, V. E., Ravichandran, K. G., Tulinsky, A., Westbrook, M., & Maraganore, J. M. (1991) Structure of the hirugen and hirulog-1 complexes of a thrombin, *J. Mol. Biol. 221*, 1379−1393.

Steiner, T., & Saenger, W. (1993) Role of C−H···O hydrogen bonds in the coordination of water molecules. Analysis of neutron diffraction data, *J. Am. Chem. Soc. 115*, 4540−4547.

Stubbs, M. T., & Bode, W. (1992) The interaction of thrombin with fibrinogen A: structural basis for its specificity, *Eur. J. Biochem. 206*, 187−195.

Stubbs, M. T., & Bode, W. (1993) A player of many parts: the spotlight falls on thrombin's structure, *Thromb. Res. 69*, 1−58.

Taylor, R., & Kennard, O. (1982) Crystallographic evidence for the existence of C−H···O, C−H···N, and C−H···Cl hydrogen bonds, *J. Am. Chem. Soc. 104*, 5063−5070.

Vitali, J., Martin, P. D., Malkowski, M. G., Robertson, W. D., Lazar, J. B., Winant, R. C., Johnson, P. H., & Edwards, B. F. P. (1992) The structure of a complex of bovine α-thrombin and recombinant hirudin at 2.8 Å resolution, *J. Biol. Chem. 267*, 17670−17678.

Vitali, J., Martin, P., Malkowski, M., Olsen, C., Johnson, P., & Edwards, B. F. P. (1996) Structure of a Bovine Thrombin:Hirudin 51−65 Complex Determined by a Combination of Molecular Replacement and Graphics. Incorporation of Known Structural Information in Molecular Replacement, *Acta Crystallogr. D52*, 453−464.

Vu, T. H. K., Hung, D. T., Wheaton, V. I., & Coughlin, S. R. (1991) Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation, *Cell 64*, 1057−1068.

Warshel, A., Naray-Szabo, G., Sussmann, F., & Hwang, J. K. (1989) How do serine proteases really work?, *Biochemistry 28*, 3629−3637.

Zdanov, A., Wu, S., DiMaio, J., Konishi, Y., Li, Y., Wu, X., Edwards, B. F. P., Martin, P. D., & Cygler, M. (1993) Crystal structure of the complex of human α-thrombin and non hydrolyzable bifunctional inhibitors hirutonin-2 and hirutonin-6, *Proteins 17*, 252−265.

BI960656Y