Protein Science (1992), *1*, 426–471. Cambridge University Press. Printed in the USA.

The refined 1.9-Å X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human α-thrombin: Structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure–function relationships

WOLFRAM BODE, DUSAN TURK,[1] AND ANDREJ KARSHIKOV
Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany

(RECEIVED October 29, 1991; REVISED MANUSCRIPT RECEIVED November 18, 1991)

Abstract

Thrombin is a multifunctional serine proteinase that plays a key role in coagulation while exhibiting several other key cellular bioregulatory functions. The X-ray crystal structure of human α-thrombin was determined in its complex with the specific thrombin inhibitor D-Phe-Pro-Arg chloromethylketone (PPACK) using Patterson search methods and a search model derived from trypsinlike proteinases of known spatial structure (Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R., & Hofsteenge, J., 1989, *EMBO J. 8*, 3467–3475). The crystallographic refinement of the PPACK–thrombin model has now been completed at an *R* value of 0.156 (8 to 1.92 Å); in particular, the amino- and the carboxy-termini of the thrombin A-chain are now defined and all side-chain atoms localized; only proline 37 was found to be in a cis-peptidyl conformation.

The thrombin B-chain exhibits the characteristic polypeptide fold of trypsinlike serine proteinases; 195 residues occupy topologically equivalent positions with residues in bovine trypsin and 190 with those in bovine chymotrypsin with a root-mean-square (r.m.s.) deviation of 0.8 Å for their α-carbon atoms. Most of the inserted residues constitute novel surface loops. A chymotrypsinogen numbering is suggested for thrombin based on the topological equivalences. The thrombin A-chain is arranged in a boomeranglike shape against the B-chain globule opposite to the active site; it resembles somewhat the propeptide of chymotrypsin(ogen) and is similarly not involved in substrate and inhibitor binding.

Thrombin possesses an exceptionally large proportion of charged residues. The negatively and positively charged residues are not distributed uniformly over the whole molecule, but are clustered to form a sandwichlike electrostatic potential; in particular, two extended patches of mainly positively charged residues occur close to the carboxy-terminal B-chain helix (forming the presumed heparin-binding site) and on the surface of loop segment 70–80 (the fibrin[ogen] secondary binding exosite), respectively; the negatively charged residues are more clustered in the ringlike region between both poles, particularly around the active site. Several of the charged residues are involved in salt bridges; most are on the surface, but 10 charged protein groups form completely buried salt bridges and clusters. These electrostatic interactions play a particularly important role in the intrachain stabilization of the A-chain, in the coherence between the A- and the B-chain, and in the surface structure of the fibrin(ogen) secondary binding exosite (loop segment 67–80).

The most remarkable feature at the thrombin surface is the prominent canyonlike active-site cleft mainly shaped by two characteristic insertion loops around Trp 60D and Trp 148. The deep and narrow active-site cleft in general explains the narrow specificity of thrombin for distinct macromolecular substrates and inhibitors. Comparisons with other crystal structures of human and bovine thrombin recently determined using this PPACK–thrombin model indicate that the first loop around Trp 60D is relatively rigid, whereas the opposite loop around Trp 148 can attain different conformations depending on complexation state and crystalline environment.

The active-site residues and the entrance to the specificity pocket are partially occluded in thrombin (much more than in the other serine proteinases) by this distinctive Trp 60D loop. The specificity pocket of thrombin resem-

Reprint requests to: Wolfram Bode, Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany.
[1] On leave from the Chemical Institut Boris Kidric, Ljubljana, Slovenia.

bles that of bovine trypsin but is designed to prefer arginine over lysine residues at P1. D-Phe 11 and Pro 21 of the bound PPACK inhibitor fit neatly to a novel hydrophobic cleft (the aryl-binding site) and to the cavitylike hydrophobic S2 subsite; the D-configuration of Phe 11 is beneficial for binding as it allows the PPACK amino-terminus to form hydrogen bonds to Gly 216 in addition. Some small arginine and benzamidine-derived synthetic inhibitors owe their particularly high thrombin specificity and affinity to their exceptional steric fit to these novel hydrophobic cavities close to the thrombin active site (Bode, W., Turk, D., & Stürzebecher, J., 1990, *Eur. J. Biochem. 193*, 175–182).

The active-site cleft levels off in the primed direction and continues over the molecular surface of the thrombin loop Lys 70–Glu 80 (itself structurally similar to the calcium loop in trypsin, with the distal nitrogen of Lys 70 replacing the calcium of trypsin). This site of a strong positive electrostatic surface potential probably represents the secondary site for interaction with the α-chain of fibrinogen and fibrin (the fibrin[ogen] secondary binding exosite) and accommodates the carboxy-terminal acidic tail part of hirudin (Rydel, T.J., Ravichandran, K.G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., & Fenton, J.W., II, 1990, *Science 249*, 277–280). Segment Arg 187–Gly 188–Asp 189, which could represent a thrombin adhesion site for cellular interactions with platelets, fibroblasts, and endothelial cells, is mainly buried in α-thrombin; its adhesive role would thus appear to require some prior unfolding.

Most of the well-characterized sites of proteolytic cleavage leading to the degradation products β-, γ-, and ε-thrombin of diminished or lost clotting activity are situated in exposed mobile loops of α-thrombin. None of these segments is in a canonical conformation that would allow association with the substrate-binding site of a cleaving serine proteinase without large conformational changes. The cleavage of the Arg 77A–Asn 78 scissile peptide bond (leading to β-thrombin) presumably results in the unfolding of the 70–80 loop, exposing the salt bridge-connected residues buried in α-thrombin to the solvent; the concomitant disruption of the surface of the fibrin(ogen) secondary binding exosite would explain the loss of binding capacity (and thus catalytic activity) toward fibrinogen. The Arg 67–Ile 68 peptide bond is completely buried beneath this 70–80 loop surface and thus only susceptible to an attacking proteinase after prior cleavage and exposure of this loop, as observed.

Keywords: electrostatic interactions; protein crystallography; serine proteinase; thrombin; thrombosis

α-Thrombin[1] is a glycosylated trypsinlike serine proteinase generated in the penultimate step of the blood coagulation cascade from the circulating plasma protein prothrombin. Upon autocatalytic and Factor Xa cleavage the functional two-chain molecule α-thrombin is generated. In the case of the human species this α-thrombin consists of the 36-residue A-chain and the 259-residue B-chain (Butkowski et al., 1977; Thompson et al., 1977; Degen et al., 1983). The two chains are connected covalently by a disulfide bridge. The B-chain has been shown to be homologous to the catalytic domains of other pancreatic and coagulation/fibrinolytic trypsinlike proteinases (Jackson & Nemerson, 1980). Upon further autolytic or proteolytic cleavage, more species (in particular β- and γ-thrombin) are generated that retain some activity against small synthetic substrates but have lost most or all clotting activity (Lundblad et al., 1979; Elion et al., 1986; Hofsteenge et al., 1988).

Thrombin is a multifunctional protein; it plays a central role in thrombosis and hemostasis but is also implicated in wound healing and various disease processes (see Fenton, 1988). α-Thrombin converts fibrinogen into fibrin, which consequently aggregates and forms the interconnecting network of thrombi. Furthermore, thrombin is able to activate several coagulation and plasma factors, such as Factors V, VIII, XIII, and protein C. The complex of thrombin with thrombomodulin exhibits enhanced reactivity toward protein C (Esmon et al., 1982) but has lost all other procoagulant activity (Esmon et al., 1983; Hofsteenge & Stone, 1987). Thrombin is effectively inhibited by only a very few endogenous protein inhibitors, such as α_2-macroglobulin and the serpins antithrombin III (Rosenberg & Damus, 1973), heparin cofactor II (Tollefsen et al., 1982), and protease nexin I (Cunningham & Farrell, 1986). The interaction with these serpins is considerably enhanced by the acidic glycosaminoglycan heparin (Li et al., 1976; Björk & Lindahl, 1982; Olson & Shore, 1982; Wallace et al., 1989). In addition, thrombin binds very tightly to and is selectively inhibited by hirudin, a protein isolated from the medicinal leech (Walsmann & Markwardt, 1981; Stone & Hofsteenge, 1986; Dodt et al., 1988).

Thrombin acts on a variety of cells (see Jackson & Nemerson, 1980). Some of these cellular interactions require proteolytically active forms of thrombin, whereas others also function with active-site-blocked thrombin species (see, e.g., Carney et al., 1986; Fenton, 1986; McGowan & Detwiler, 1986; Shuman, 1986; Vu et al., 1991). Thrombin induces platelet aggregation and stim-

[1] P1, P2, P3, etc., and P1′, P2′ designate substrate/inhibitor residues amino- and carboxy-terminal to the scissile peptide bond, respectively, and S1, S2, S3, etc., and S1′, S2′ the corresponding subsites of the cognate proteinase (Schechter & Berger, 1967). For numbering of the thrombin amino acid residues, the chymotrypsinogen nomenclature introduced by Bode et al. (1989b) is used. D-Phe-Pro-Arg chloromethylketone (PPACK) residues are designated by a suffix I, hirudin residues by a suffix HI behind the residue number.

Deposition: The PPACK-thrombin have been deposited at the Brookhaven Protein Data Bank (Bernstein et al., 1977).

ulates platelet secretion. It causes mitogenesis in fibroblasts (Cunningham & Farrell, 1986; Glenn et al., 1988) and macrophagelike cells (Bar-Shavit et al., 1986), exhibits chemotactic properties (Bar-Shavit et al., 1983, 1984, 1986; Bizios et al., 1986), and binds to endothelial cells (Prescott et al., 1990; Bar-Shavit et al., 1991) and to subendothelial extracellular matrix (Bar-Shavit et al., 1989).

Thrombin cleaves a variety of proteins for which the amino acid sequence around the scissile peptide bond differs considerably (Blombäck et al., 1977; Scheraga, 1977). Its specificity is primarily trypsinlike, i.e., it cleaves behind arginine and lysine residues, with a clear preference for Arg-Xaa bonds (Liem & Scheraga, 1974; Lottenberg et al., 1983). It is, however, much more selective than trypsin toward macromolecular substrates in that it cleaves many fewer peptide bonds of permitted sequence; in fibrinogen, for example, only two Arg-Xaa bonds out of a total of 181 Arg/Lys-Xaa bonds are attacked (Blombäck et al., 1967; Hogg & Blombäck, 1978). Thrombin is particularly active toward peptidic substrates and inhibitors with a proline at the P2 position (Kettner & Shaw, 1981); a D-phenylalanine residue at P3 makes such peptides even more reactive toward thrombin. Rather reactive and selective chloromethyl, aldehyde, and boroarginine inhibitors have been developed with the peptidyl moiety D-Phe-Pro-Arg (Bajusz et al., 1978; Kettner & Shaw, 1979; Kettner et al., 1991); various synthetic thrombin inhibitors have been found and characterized based on arginine and benzamidine derivatives or on heterocyclic compounds (Okamoto et al., 1981; Stürzebecher et al., 1983; Kam et al., 1988).

In contrast to most other serine proteinases, the thrombin specificity toward fibrinogen is not determined by subsites surrounding the active residues alone. For efficient cleavage of fibrinogen (as well as of isolated Aα-chains) the availability and integrity of a thrombin exosite quite distant from the catalytic residues is important (Hageman & Scheraga, 1974; Liem & Scheraga, 1974; Blombäck et al., 1977; Van Nispen et al., 1977; Marsh et al., 1985). This fibrin(ogen)-recognizing exosite, attributed to a rather basic surface area of thrombin around its β-cleavage site has also been implicated in binding of other proteins such as fibrin monomers, fibrin E-domain, and its NDSK-fragment, thrombomodulin, and hirudin, and of negatively charged cell surfaces, glass, and cation-exchange resins (Liu et al., 1979; Fenton et al., 1981, 1988; Berliner et al., 1985; Kaminski & McDonagh, 1987; Kaczmarek & McDonagh, 1988; Vali & Scheraga, 1988).

An apolar binding site close to the catalytic center has been inferred from binding and proflavine displacement studies (Thompson, 1976; Berliner & Shen, 1977). This site has been suggested as accounting for thrombin specificity with tripeptide substrates (Sonder & Fenton, 1984) and for the accommodation of large hydrophobic residues amino-terminal of the fibrinogen Aα-cleavage site (Meinwald et al., 1980; Marsh et al., 1985; Ni et al., 1989a,b).

Several dysfunctional genetic variants of human (pro)thrombin have been found; in three of them the site of mutation has been localized (Miyata et al., 1987; Henriksen & Mann, 1988, 1989). Recently, several recombinant thrombin mutants also have been prepared (LeBonniec & Esmon, 1991; LeBonniec et al., 1991; Wu et al., 1991).

Three-dimensional models have been proposed for the thrombin B-chain (Magnusson et al., 1975; Greer, 1981; Bing et al., 1986; Sugawara et al., 1986; Toma & Suzuki, 1989) based on the crystal structures of bovine trypsin and chymotrypsin. These models provided a general impression of the arrangement of sites involved in the various interactions of thrombin and were of some use in explaining certain structure–function relationships and in developing inhibitors (Maraganore et al., 1990). An adequate understanding of the specific and characteristic functions of thrombin and the accurate model-based design of thrombin inhibitors requires, however, knowledge of an experimentally determined structure.

We have recently communicated the X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone (PPACK [Kettner & Shaw, 1979]) inhibited human α-thrombin and presented a concise description of the most important features of thrombin, in particular its unique and prominent canyonlike active-site cleft (Bode et al., 1989b). This PPACK–thrombin model is now fully refined; the relatively flexible amino- and carboxy-termini of the A-chain have been determined, the locations of a few amino acid side chains have been revised, and more solvent molecules have been added in order to improve further the phases.

Using the refined coordinates of this PPACK–thrombin structure as a search and initial phasing model, the X-ray crystal structures of human α-thrombin complexes of hirudin (Grütter et al., 1990; Rydel et al., 1990, 1991) and fibrinopeptide A/hirugen (Stubbs et al., 1992) and of bovine thrombin complexes (Martin et al., 1992) have been elucidated. Recently, we established the exact binding geometry of some arginine- and benzamidine-based synthetic inhibitors toward bovine thrombin (Brandstetter et al., 1992) and confirmed earlier modeling results inferred from trypsin binding (Bode et al., 1990; Chow et al., 1990; Turk et al., 1991).

In the following, the entire course of the structure analysis of PPACK–thrombin and the completely refined model are fully described; sites of particular interest, such as the A–B-chain interaction and the distinct substrate-binding subsites including the interactions with substrates and inhibitors are presented in more detail; the unique electrostatic properties of thrombin that play an important role in the molecular stability and in several molecular functions are described; the thrombin structure is compared with that of related proteinases; and finally, the various functions of thrombin are discussed with respect to its structure.

Results and discussion

Overall structure

The thrombin molecule can be described as a prolate ellipsoid of approximate dimensions 45 Å × 45 Å × 50 Å. Its A- and B-chains are not organized in separate domains, but form a single contiguous body with a highly furrowed surface (see, e.g., Fig. 1). Similar to the other trypsinlike serine proteinases, thrombin consists essentially of two interacting six-stranded barrellike domains (linked by four transdomain "straps"), of five helical segments and one helical turn, and of various surface-located turn structures (Fig. 2; Kinemage 1). Three of these straps and three of these helices are also present in other known serine proteinase structures; the A-chain has a counterpart in chymotrypsin(ogen). The catalytic residues, in particular Ser 195, His 57, and Asp 102 (using the chymotrypsinogen nomenclature introduced by Bode et al. [1989b]; see Table 3) are located at the junction between both barrels; a prominent active-site cleft (Fig. 1) stretches perpendicular to this junction (described in more detail below).

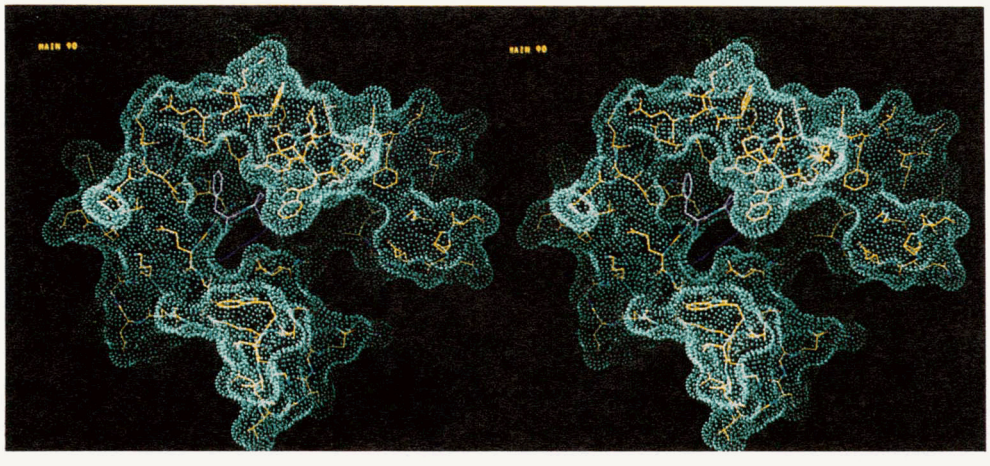

Fig. 1. Front view toward the thrombin molecule (yellow), displayed together with its Connolly dot surface (calculated with a probe of radius 1.4 Å). The active-site cleft runs from left to right across the molecular surface; the bound PPACK molecule (violet) is bound. Only the front parts of the thrombin molecule and of its molecular surface are displayed. A bound substrate polypeptide chain would run from left to right. The surface "hole" close to the center corresponds to the entrance to the specificity pocket. The 60 insertion loop, in particular Trp 60D, is partially occluding the active site. The fibrinogen secondary binding exosite is to the right.

Fig. 2. Ribbon plot of the human α-thrombin polypeptide chain; β-strands, helices, and turns are represented by twisted arrows, helical ribbons, and ropes. The view is on the active-site cleft made by the B-chain (yellow); the A-chain runs in the back of the molecule (violet). This figure has been produced with a modified version of the program RIBBON kindly provided by John Priestle (Priestle, 1988).

The secondary structure of the A- and the B-chain of the final thrombin model is illustrated schematically in Figure 3. Each barrellike domain appears as six antiparallel running strands. The size and extent of each β-sheet is remarkably similar to those observed in other trypsin-like serine proteinases. The polypeptide strands involved in barrel formation always fold back on the molecular surface (see Fig. 2); some of these surface-located thrombin loops are particularly extended in comparison to other related serine proteinases. In each barrel, four of the five turns are hairpin loops, all of them of a rather open, nonregular conformation. Some of these open loops contain local classical 1–4 tight turns (Table 1).

The six helical segments occurring in thrombin are listed in Table 2 together with their helix type and length defined according to different helix criteria. The one-turn helix around His 57, the intermediate helix around Cys 168, and the long carboxy-terminal helix (of a mixed type) are similar to those found in other serine proteinases; the A-chain helix, the short 3_{10}-helix between Tyr 60A and Asp 60E (better described as two interlaced type-III turns; see Table 1), and the regular two-turn α-helix around Ala 129 are additional segments or parts of insertion loops and seem to be characteristic of thrombin.

Comparison with related serine proteinase structures

The α-carbon structure of human α-thrombin optimally superimposed on that of bovine chymotrypsinogen is

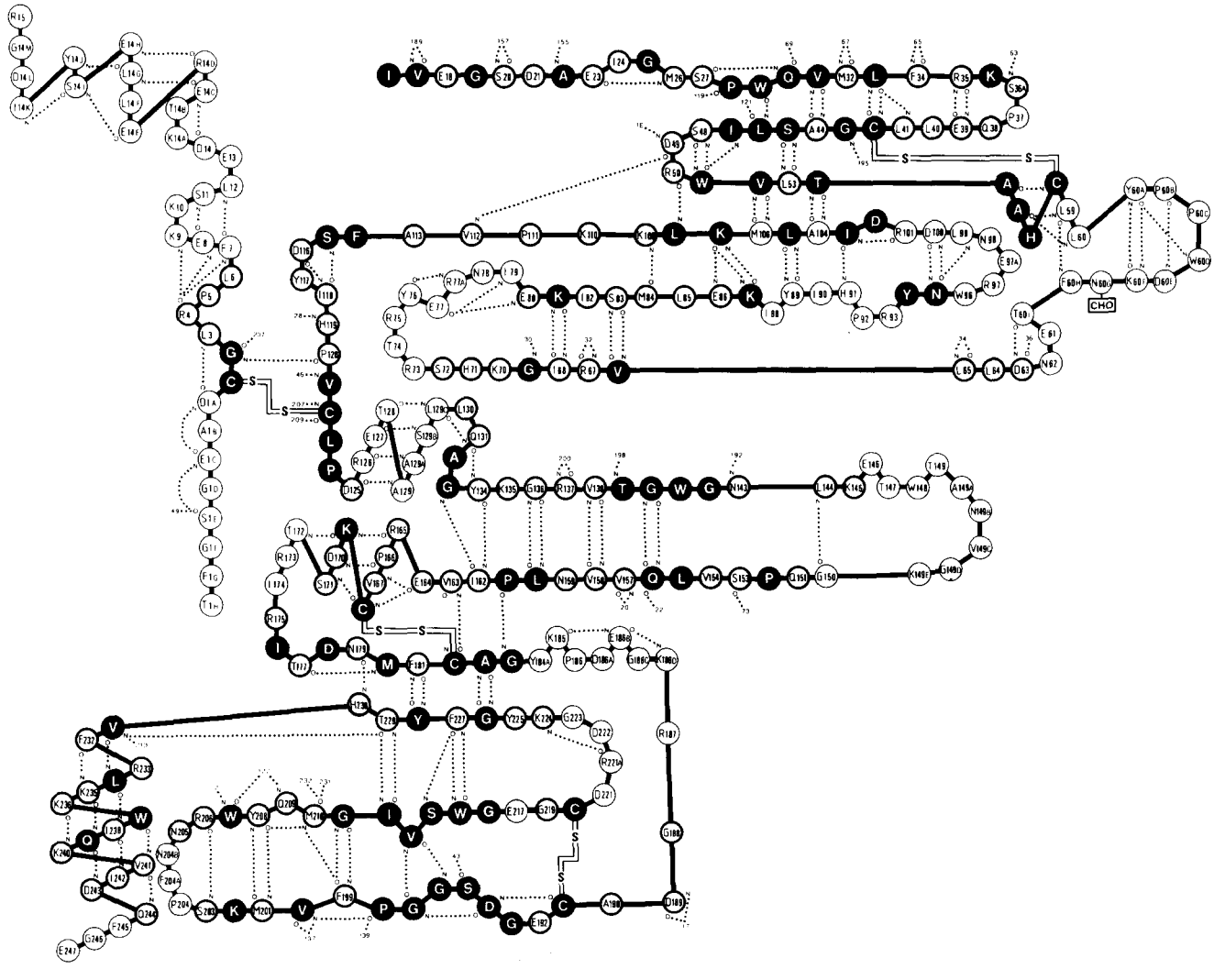

Fig. 3. Secondary structure of human α-thrombin. Inter-main-chain hydrogen bonds are selected according to criteria (with $E < -0.7$ kcal/mol) given by Kabsch and Sander (1983). The sequence nomenclature is that derived from equivalency with chymotrypsin(ogen)/trypsin (Bode et al., 1989b) as also used in Table 3. Thrombin residues that are topologically equivalent to chymotrypsin residues are emphasized by thick circles; thrombin residues, which in addition possess identical amino-acids, are denoted by filled circles.

Table 1. *Tight 1-4 turns in human α-thrombin*

Residues	Position 1	2	3	4	Turn type	Hydrogen bond: 1-4 Hydrogen-bond length (Å)	Energy (kcal/mol)[a]	Remarks
1A-3	Asp	Cys	Gly	Leu	II	3.45	−1.3	Gly in position 3
4-7	Arg	Pro	Leu	Phe	I-III	3.01	−1.5	Pro in position 2
8-11	Glu	Lys	Lys	Ser	I	2.89	−1.9	
14-14C	Asp	Lys	Thr	Glu	I	3.21	−1.6	
14G-14J	Leu	Glu	Ser	Tyr	III	2.80	−2.1	
23-26	Glu	Ile	Gly	Met	II	3.40	−1.3	Gly in position 3
48-51	Ser	Asp	Arg	Trp	I	3.46	−0.7	Very weak hydrogen bond
60A-60D	Tyr	Pro	Pro	Trp	III	3.20	−1.5	Part of 3_{10}-helix
60B-60E	Pro	Pro	Trp	Asp	I	2.74	−2.8	Pro not in position 3
60I-63	Thr	Glu	Asn	Asp	I	2.97	−1.8	
77-79	Glu	Arg	Asn	Ile	I	3.41	−0.8	Very weak hydrogen bond
91-94	His	Pro	Arg	Tyr	I	3.39	−1.0	Pro not in position 3
131-134	Gln	Ala	Gly	Tyr	II	2.84	−2.4	Gly in position 3
164-167	Glu	Arg	Pro	Val	III	3.07	−0.9	
168-171	Cys	Lys	Asp	Ser	I	3.11	−1.6	
177-180	Thr	Asp	Asn	Met	I	3.12	−1.5	
185-186B	Lys	Pro	Asp	Glu	I	3.01	−2.2	Pro not in position 3
191-194	Cys	Glu	Gly	Asp	II	3.34	−1.5	Gly in position 3
221-224	Arg	Asp	Gly	Lys	II	3.18	−1.7	Gly in position 3

[a] According to Kabsch and Sander (1983).

shown in Figure 4. One hundred ninety amino acid residues of the thrombin B-chain are topologically equivalent (within a root-mean-square [r.m.s.] deviation of 0.78 Å) to residues in bovine chymotrypsin (Cohen et al., 1981; Blevins & Tulinsky, 1985; Tsukada & Blow, 1985); in addition, 6 residues of the thrombin A-chain are equivalent to the 6 amino-terminal residues of the chymotrypsin(ogen) propeptide. The 190 equivalent B-chain residues and another 40 residues were assigned the sequence numbers of the topologically equivalent chymotrypsinogen residues (Hartley & Kauffman, 1966; Meloun et al., 1966). In addition to secondary structural elements, Figure 3 defines those human α-thrombin amino acid residues topologically equivalent to chymotrypsin residues (see also Kinemage 2); filled circles further represent those 84 with identical amino acids.

The thrombin B-chain has 195 α-carbon atoms equivalent to the bovine trypsin structure (see Fig. 5) with an r.m.s. deviation of 0.8 Å (Bode & Schwager, 1975a); at a few sites the thrombin structure is closer to that of trypsin than to that of chymotrypsin, i.e., (1) thrombin segment 70–80, which resembles the calcium-binding loop of trypsin (Bode & Schwager, 1975a,b), (2) Gly 184–Tyr 184A, and (3) thrombin loop segment 217–224.

Here, the generally used chymotrypsinogen numbering of bovine trypsin (Hartley & Shotton, 1971) was used.

Table 2. *Helical segments in human α-thrombin*

Helix type	Helix extent (number of residues) according to Richardson and Richardson (1990): Hydrogen bonding	Conformation	α-Carbon position
3.6_{13}	Glu 14C to Ser 14I (7)	Glu 14C to Ser 14I (7)	Thr 14B to Tyr 14J (9)
3_{10}	Ala 55 to Leu 58 (5)	Ala 56 to Leu 59 (5)	Ala 56 to Tyr 60A (6)
3_{10}	Tyr 60A to Asp 60E (5)	Pro 60B to Pro 60C (2)	Tyr 60A to Trp 60D (4)
3.6_{13}	Asp 125 to Leu 129C (8)	Arg 126 to Ser 129B (6)	Asp 125 to Glu 131 (10)
3.6_{13}	Glu 164 to Asp 170 (7)	Arg 165 to Lys 169 (5)	Glu 164 to Ser 171 (8)
3_{10}	Cys 168 to Thr 172 (5)	Lys 169 to Ser 171 (3)	Lys 169 to Thr 172 (4)
3_{10}	Val 231 to Lys 235 (5)	Val 231 to Trp 237 (7)	His 230 to Leu 234 (5)
3.6_{13}	Leu 234 to Asp 243 (10)	Ile 238 to Ile 242 (5)	Leu 234 to Ile 242 (9)
3_{10}	Val 241 to Glu 244 (4)	Val 241 to Glu 244 (4)	Val 241 to Glu 244 (4)

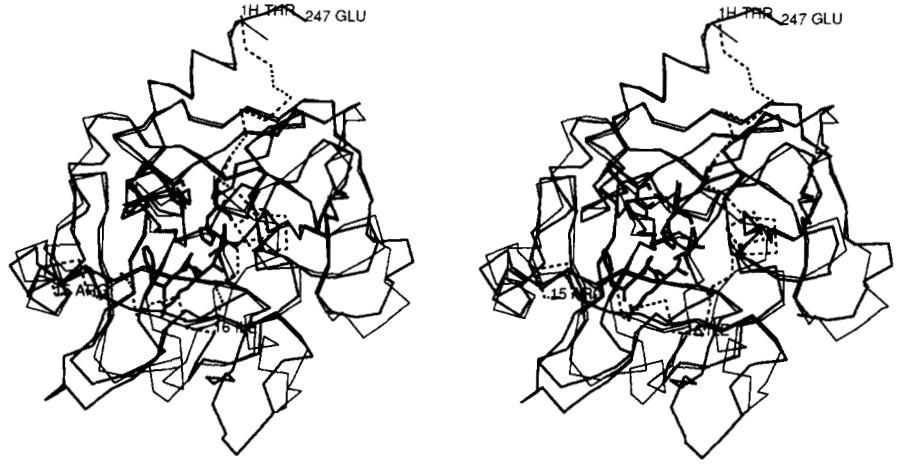

Fig. 4. α-carbon structure of the A-chain (dashed connections) and the B-chain (thick connections) of human α-thrombin superimposed with molecule 1 of bovine chymotrypsinogen (thin connections [Wang et al., 1985]) after minimizing the r.m.s. deviation of the topologically equivalent α-carbon atoms. The PPACK molecule is shown with all nonhydrogen atoms, and the thrombin chain termini are labeled.

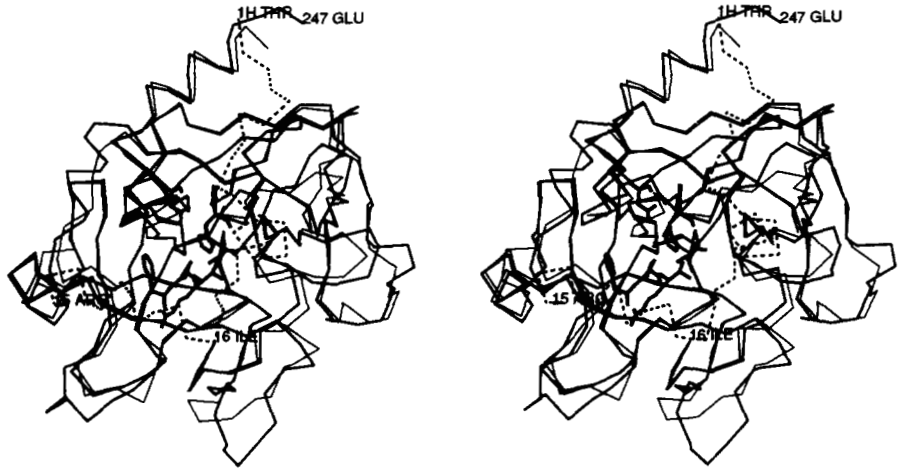

Fig. 5. α-carbon structure of the A-chain (dashed connections) and the B-chain of human α-thrombin (thick connections) superimposed with bovine trypsin (Bode et al., 1976). The PPACK molecule is overlaid; standard view as in Figure 1.

The residual insertion residues in the thrombin B-chain were marked by letter suffixes; at a few of the junctions between conserved and variable segments, the positioning of the insertions was necessarily somewhat arbitrary.

For numbering of the A-chain, residue numbers 1 and 15 were assigned to the (chymotrypsin-homologous) cysteine and to the carboxy-terminal arginine residue, respectively. The A-chain residues preceding Cys 1 were designated by a 1 and a letter suffix in alphabetic order in the reverse chain direction; the A-chain residues following Cys 1 were designated by numbers up to 14, followed by 14 and letter suffixes in alphabetic order.

The chymotrypsinogen sequence numbering of human thrombin is shown in Table 3 based on these topological equivalences with chymotrypsin and trypsin (see Bode et al., 1989b). The A-chain runs from Thr 1H to Arg 15; the B-chain starts with Ile 16 and ends up with Glu 247. A nomenclature based on the topological equivalence of the amino acid residues is an enormous advantage, as functional properties already known for the related trypsinlike enzymes can be attributed to distinct sites of the thrombin structure. Table 3 also contains the sequences of bovine chymotrypsinogen (Hartley & Kauffman, 1966; Meloun et al., 1966), bovine trypsinogen (Walsh & Neurath, 1964; Mikes et al., 1966), and bovine thrombin (Magnusson et al., 1975; MacGillivray & Davie, 1984). All key residues of thrombin are also conserved in the corresponding enzymes from mouse (Friezner-Degen et al., 1990) and rat (Dihanich & Monard, 1990).

According to this topological alignment, 86 (of the equivalent) amino acid residues of the thrombin B-chain are identical with residues in bovine trypsin, and 79 are identical with residues in bovine chymotrypsin (whose propeptide possesses, in addition, two residues identical to the thrombin A-chain). For comparison, the two digestive enzymes have 99 identical residues between them.

Of the other vertebrate serine proteinases of known spatial structure, porcine pancreatic kallikrein (Bode et al., 1983) and porcine pancreatic elastase (Meyer et al., 1988) exhibit a similar topological equivalence to the thrombin B-chain; however, their r.m.s. deviation is slightly larger and the number of identical residues slightly smaller (Ta-

Table 3. *Sequence alignment of bovine thrombin (BTHR), human thrombin (HTHR), bovine trypsin (BTRY), and bovine chymotrypsin (BCHY) according to topological equivalences; thrombin chymotrypsinogen numbering (THRO), chymotrypsinogen numbering (CHYG)*[a]

THRO	1H	1G	1F	1E	1D	1C	1B	1A	1	2	3	4	5	6	7
BTHR	T	F	G	A	G	E	A	D	C	G	L	R	P	L	F
HTHR	T	F	G	S	G	E	A	D	C	G	L	R	P	L	F
BTRY															
BCHY									C	G	V	P	A	I	Q
CHYG									1	2	3	4	5	6	7

THRO	8	9	10	11	12	13	14	14A	14B	14C	14D	14E	14F	14G	14H
BTHR	E	K	K	Q	V	Q	D	Q	T	E	K	E	L	F	E
HTHR	E	K	K	S	L	E	D	K	T	E	R	E	L	L	E
BTRY															
BCHY	P	V	L	S	G	L	S	-	-	-	-	-	-	-	-
CHYG	8	9	10	11	12	13	14	-	-	-	-	-	-	-	-

THRO	14I	14J	14K	14L	14M	15
BTHR	S	Y	I	E	G	R
HTHR	S	Y	I	D	G	R
BTRY	V	D	D	D	D	K
BCHY	-	-	-	-	-	R
CHYG	-	-	-	-	-	15

THRO	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
BTHR	I	V	E	G	Q	D	A	E	V	G	L	S	P	W	Q
HTHR	I	V	E	G	S	D	A	E	I	G	M	S	P	W	Q
BTRY	I	V	G	G	Y	T	C	G	A	N	T	V	P	Y	Q
BCHY	I	V	N	G	E	E	A	V	P	G	S	W	P	W	Q
CHYG	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

THRO	31	32	33	34	35	36	36A	37	38	39	40	41	42	43	44
BTHR	V	M	L	F	R	K	S	CP	Q	E	L	L	C	G	A
HTHR	V	M	L	F	R	K	S	CP	Q	E	L	L	C	G	A
BTRY	V	S	L	N	S	-	-	-	G	Y	H	F	C	G	G
BCHY	V	S	L	Q	D	K	-	T	G	F	H	F	C	G	G
CHYG	31	32	33	34	35	36	-	37	38	39	40	41	42	43	44

THRO	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
BTHR	S	L	I	S	D	R	W	V	L	T	A	A	H	C	L
HTHR	S	L	I	S	D	R	W	V	L	T	A	A	H	C	L
BTRY	S	L	I	N	S	Q	W	V	V	S	A	A	H	C	Y
BCHY	S	L	I	N	E	N	W	V	V	T	A	A	H	C	G
CHYG	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59

THRO	60	60A	60B	60C	60D	60E	60F	60G	60H	60I	61	62	63	64	65
BTHR	L	Y	P	P	W	D	K	N	H	I	V	D	D	L	L
HTHR	L	Y	P	P	W	D	K	N	H	I	E	N	D	L	L
BTRY	K	-	-	-	-	-	-	-	-	-	-	S	G	I	Q
BCHY	V	-	-	-	-	-	-	-	-	-	T	T	S	D	V
CHYG	-	-	-	-	-	-	-	-	-	-	61	62	63	64	65

THRO	66	67	68	69	70	71	72	73	74	75	76	77	77A	78	79
BTHR	V	R	I	G	K	H	S	R	T	R	Y	E	R	K	V
HTHR	V	R	I	G	K	H	S	R	T	R	Y	E	R	N	I
BTRY	V	R	L	G	E	D	N	I	N	V	V	E	-	G	N
BCHY	V	V	A	G	E	F	D	Q	G	S	S	S	-	E	K
CHYG	66	67	68	69	70	71	72	73	74	75	76	77	77A	78	79

THRO	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94
BTHR	E	K	I	S	M	L	D	K	I	Y	I	H	P	R	Y
HTHR	E	K	I	S	M	L	E	K	I	Y	I	H	P	R	Y
BTRY	E	Q	F	I	S	A	S	K	S	I	V	H	P	S	Y
BCHY	I	Q	K	L	K	I	A	K	V	F	K	N	S	K	Y
CHYG	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94

(*continued*)

Table 3. *Continued*

THRO	95	96	97	97A	98	99	100	101	102	103	104	105	106	107	108
BTHR	N	W	K	E	N	L	D	R	D	I	A	L	L	K	L
HTHR	N	W	R	E	N	L	D	R	D	I	A	L	M	K	L
BTRY	N	S	N	-	T	L	N	N	D	I	M	L	I	K	L
BCHY	N	S	L	-	T	I	N	N	D	I	T	L	L	K	L
CHYG	95	96	97	-	98	99	100	101	102	103	104	105	106	107	108

THRO	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123
BTHR	K	R	P	I	E	L	S	D	Y	I	H	P	V	C	L
HTHR	K	K	P	V	A	F	S	D	Y	I	H	P	V	C	L
BTRY	K	S	A	A	S	L	N	S	R	V	A	S	I	S	L
BCHY	S	T	A	A	S	F	S	Q	T	V	S	A	V	C	L
CHYG	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123

THRO	124	125	126	127	128	129	129A	129B	129C	130	131	132	133	134	135
BTHR	P	D	K	Q	T	A	A	K	L	L	H	A	G	F	K
HTHR	P	D	R	E	T	A	A	S	L	L	Q	A	G	Y	K
BTRY	P	T	-	S	C	A	-	-	-	-	S	A	G	T	Q
BCHY	P	S	A	S	D	D	-	-	-	F	A	A	G	T	T
CHYG	124	125	126	127	128	129	-	-	-	130	131	132	133	134	135

THRO	136	137	138	139	140	141	142	143	144	145	146	147	148	149	149A
BTHR	G	R	V	T	G	W	G	N	R	R	E	T	W	T	T
HTHR	G	R	V	T	G	W	G	N	L	K	E	T	W	T	A
BTRY	C	L	I	S	G	W	G	N	T	K	S	S	G	T	-
BCHY	C	V	T	T	G	W	G	L	T	R	Y	T	N	A	-
CHYG	136	137	138	139	140	141	142	143	144	145	146	147	148	149	-

THRO	149B	149C	149D	149E	150	151	152	153	154	155	156	157	158	159	160
BTHR	S	V	A	E	V	Q	P	S	V	L	Q	V	V	N	L
HTHR	N	V	G	K	G	Q	P	S	V	L	Q	V	V	N	L
BTRY	-	-	-	-	S	Y	P	D	V	L	K	C	L	K	A
BCHY	-	-	-	-	N	T	P	D	R	L	Q	Q	A	S	L
CHYG	-	-	-	-	150	151	152	153	154	155	156	157	158	159	160

THRO	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
BTHR	P	L	V	E	R	P	V	C	K	A	S	T	R	I	R
HTHR	P	I	V	E	R	P	V	C	K	D	S	T	R	I	R
BTRY	P	I	L	S	D	S	S	C	K	S	A	Y	P	G	Q
BCHY	P	L	L	S	N	T	N	C	K	K	Y	W	G	T	K
CHYG	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175

THRO	176	177	178	179	180	181	182	183	184	184A	185	186	186A	186B	186C
BTHR	I	T	D	N	M	F	C	A	G	Y	K	P	G	E	G
HTHR	I	T	D	N	M	F	C	A	G	Y	K	P	D	E	G
BTRY	I	T	S	N	M	F	C	A	G	Y	L	E	-	-	-
BCHY	I	K	D	A	M	I	C	A	G	-	A	S	-	-	-
CHYG	176	177	178	179	180	181	182	183	184	-	185	186	-	-	-

THRO	186D	187	188	189	190	191	192	193	194	195	196	197	198	199	200
BTHR	K	R	G	D	A	C	E	G	D	S	G	G	P	F	V
HTHR	K	R	G	D	A	C	E	G	D	S	G	G	P	F	V
BTRY	G	G	K	D	S	C	Q	G	D	S	G	G	P	V	V
BCHY	-	G	V	S	S	C	M	G	D	S	G	G	P	L	V
CHYG	-	187	188	189	190	191	192	193	194	195	196	197	198	199	200

THRO	201	202	203	204	204A	204B	205	206	207	208	209	210	211	212	213
BTHR	M	K	S	P	Y	N	N	R	W	Y	Q	M	G	I	V
HTHR	M	K	S	P	F	N	N	R	W	Y	Q	M	G	I	V
BTRY	C	S	-	-	-	-	-	-	G	K	L	Q	G	I	V
BCHY	C	K	K	N	-	-	G	A	W	T	L	V	G	I	V
CHYG	201	202	203	204	-	-	205	206	207	208	209	210	211	212	213

(continued)

Table 3. *Continued*

THRO	214	215	216	217	-	219	220	221	221A	222	223	224	225	226	227
BTHR	S	W	G	E	-	G	C	D	R	D	G	K	Y	G	F
HTHR	S	W	G	E	-	G	C	D	R	D	G	K	Y	G	F
BTRY	S	W	G	S	-	G	C	A	Q	K	N	K	P	G	V
BCHY	S	W	G	S	S	T	C	S	-	T	S	T	P	G	V
CHYG	214	215	216	217	218	219	220	221	-	222	223	224	225	226	227

THRO	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242
BTHR	Y	T	H	V	F	R	L	K	K	W	I	Q	K	V	I
HTHR	Y	T	H	V	F	R	L	K	K	W	I	Q	K	V	I
BTRY	Y	T	K	V	C	N	Y	V	S	W	I	K	Q	T	I
BCHY	Y	A	R	V	T	A	L	V	N	W	V	Q	Q	T	L
CHYG	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242

THRO	243	244	245	246	247
BTHR	D	R	L	G	S
HTHR	D	Q	F	G	E
BTRY	A	S	N	-	-
BCHY	A	A	N	-	-
CHYG	243	244	245	246	247

[a] Residues topologically equivalent with human thrombin are boxed.

ble 4). The corresponding numbers (Table 4) for rat mast cell proteinase II (Remington et al., 1988) and human leukocyte elastase (Bode et al., 1986), two cellular serine proteinases, indicate a more distant structural relationship to thrombin.

This ranking order due to topological equivalence agrees approximately with the order obtained from a sequence alignment of these serine proteinases with the thrombin B-chain (Table 4). Table 4 shows furthermore the alignment score and the degree of similarity to the thrombin B-chain for the reactive domains of several coagulation factors based on sequence alignment alone. According to such data, the human thrombin B-chain is most similar to the catalytic domains of protein C and Factor Xa of human origin; its similarity with coagulation Factors IXa and XIa is only slightly lower, in agree-

Table 4. *Topological equivalence and sequence identity of human thrombin B-chain with the reactive domains of vertebrate serine proteinases. Sequence alignment scores for other coagulation factors*[a]

	Topologically equivalent residues		Identical residues (from sequence alignment)		
	Number	r.m.s. deviation (Å)	Number	% Total (enzyme)	Alignment score (standard deviations)
BTRY	195	0.82	88	39.5	27
BCHY	190	0.78	85	37.0	24
PPKK	193	0.92	70	30.2	19
PPE	192	1.02	62	25.8	12
RMCP	180	0.89	56	25.0	13
HLE	176	0.90	61	28.0	12
(BTRY-BCHY)	190	0.76	96	43.0	35
Protein C			100	40.0	41
FXa (reactive domain)			97	38.2	40
FIXa (reactive domain)			87	37.0	37
FXIa (reactive domain)			88	37.0	34

[a] BTRY, bovine trypsin (Bode & Schwager, 1975a); BCHY, bovine chymotrypsin A (Cohen et al., 1981); PPKK, porcine pancreatic kallikrein (Bode et al., 1983); RMCP, rat mast cell proteinase II (Remington et al., 1988); PPE, porcine pancreatic elastase (Meyer et al., 1988); HLE, human leukocyte elastase (Bode et al., 1986).

ment with previous investigations of others (see, e.g., Patthy, 1985).

The polypeptide alignment (Table 3) obtained for thrombin and chymotrypsin(ogen) due to topological equivalence is in good agreement with the alignment scheme proposed by Hartley and Shotton (1971). The topologically equivalent regions of the thrombin B-chain (Table 3) contain the (structurally) conserved regions previously identified by Greer (1981) and Furie et al. (1982) on the basis of amino acid sequence homology or from comparison of the three-dimensional models of serine proteinases, respectively. The sites and numbers of insertions and deletions predicted by Bing et al. (1986) on the basis of a chymotrypsin–trypsin hybrid model are largely in accord with our experimental results. Their proposed alignment turned out to be incorrect only around Glu 80; as shown below, a single insertion at Arg 77A is found compared with the predicted deletion and insertion at positions 76 and 84, respectively. Some spatial properties of thrombin such as the positively charged anion-binding exosite and the clustering of aromatic residues near the active site were correctly predicted (Bing et al., 1986). It had been suggested that one or both large insertion loops might fold (partially) over the entrance to the catalytic center, in this way blocking the approach of large macromolecular substrates or inhibitors (Sugawara et al., 1986). Not foreseen was the role of Ile 174 (which is not part of an insertion segment, see below) or some other structural details largely responsible for thrombin specificity.

Comparison with related thrombin structures

The availability of different crystal structures of thrombin from different species allows us to recognize intrinsic structural properties and to separate them from crystal-packing effects. Comparison of the α-carbon tracing of PPACK-thrombin with that of the hirudin-thrombin complex (Rydel et al., 1991; see Fig. 6) shows both thrombin models to be to a large extent identical; two-thirds of all α-carbon atoms (i.e., 200 largely internal residues) differ by only 0.26 Å (r.m.s. deviation). There are, however, some very significant displacements and conformational differences, in particular in the amino-terminal segment (up to Ala 1B) and in the carboxy-terminal segment (from Ile 14K onward) of the A-chain, in the insertion loop around Trp 148, and at the carboxy-terminus of the B-chain from Gln 244 onward.

The displacements of the residues around Trp 148 are the result of a large conformational change of the corresponding loop, mainly through a rigid-body rotation of the loop domain around Glu 146 Cα-C and Gly 150 N-Cα (see Fig. 17 and below). The central residues of this loop are displaced by up to 10 Å (Rydel et al., 1991). Without rearrangements, the hirudin chain as it is observed in the hirudin–thrombin complex would collide with the Trp 148 loop in its PPACK–thrombin conformation (in particular, the side chains of Trp 148 and Thr 4IH [see Fig. 17]). It should be mentioned that this loop segment in both thrombin structures exhibits enhanced, but by no means extreme, B-values (see Fig. 28).

Significant deviations (up to almost 1.0 Å) are furthermore observed at the central residues of the insertion loop around Trp 60D (which juxtaposes the Trp 148 loop across the active-site cleft). These displacements result, however, from expansion of the S2 cavity (see below) enclosed by this exposed loop rather than from conformational changes. Again, the B-values of this loop are not exceptionally high in either structure.

Further significant, but more local, positional changes between PPACK–thrombin and hirudin–thrombin are observed at Glu 97A (which forms the ceiling of the hydrophobic pocket) and at Arg 77A (in the fibrin[ogen] secondary binding exosite); both residues are parts of quite exposed loops in the substrate-binding area and are rather flexible; thus their conformation might easily be affected by the differently bound inhibitors.

Comparison of the human PPACK–thrombin model with the two independent molecules (Martin, Kunjummen, Kumar, Bode, Huber, & Edwards, unpubl.) in the bovine thrombin structure (with 87% identical residues

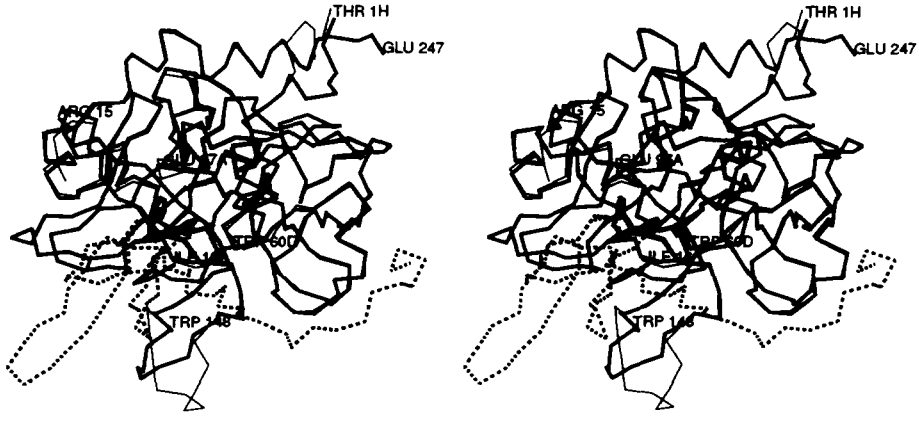

Fig. 6. α-carbon structure of PPACK–human α-thrombin (thick connections) optimally superimposed with the thrombin (thin connections) and the hirudin component (dashed connections) of the thrombin–hirudin complex (Rydel et al., 1991). Standard view as in Figure 1.

[compare with Table 3]) yields a similar pattern of polypeptide chain displacements: the free amino- and the free carboxy-terminal end of the A-chain, the carboxy-terminus of the B-chain, and the Trp 148 loop (which is only weakly defined presumably due to partial cleavage) show very large displacements, mainly as the result of different crystal contacts. Omission of these segments (i.e., of 21 amino acid residues) yields an r.m.s. deviation of only 0.45 Å for α-carbon atoms. It is noteworthy that the 60-insertion loop is similarly folded in the two bovine thrombin molecules as in PPACK–thrombin.

Charged groups and salt bridges

Human thrombin has a particularly high proportion of polar amino acids. Only about 45% of its total residues are nonpolar (compared with, for example, 51% in bovine chymotrypsin and 63% in the much more "sticky" human leukocyte elastase); more than 70% of all residues on the thrombin surface are hydrophilic in nature. Moreover, human α-thrombin possesses an exceptionally large number of charged residues compared with the pancreatic trypsinlike enzymes. The A- and the B-chain contain 9 and 31 acidic residues and 6 and 37 basic residues (and an additional 4 histidines), respectively (Fig. 7). In contrast, bovine trypsinogen and bovine chymotrypsinogen contain only 12 or 14 acidic, and 17 or 18 basic residues. In addition, human α-thrombin has two terminal amino, two terminal carboxy, and two neuraminic acid groups, the latter terminating the single Asn 60G-linked sugar chain of thrombin (Nilsson et al., 1983).

The calculated charges and electrostatic energies of all (partially) charged protein groups in the electrostatic multipole field created by the whole thrombin molecule are listed in Table 5. About three-quarters of all charged groups exhibit negative electrostatic interaction energies, i.e., confer overall stability to the thrombin molecule, whereas one-quarter of them have a destabilizing influence on the molecular structure. At pH 7, His 57 should carry an almost full positive charge (due to the close proximity of Asp 102); His 119 should be weakly charged at this pH value, whereas the residual two histidine residues, His 71 and His 91, should virtually be noncharged.

The almost balanced number of acidic and basic residues of human thrombin suggests that its overall net charge (which we calculate to be +2.3 at pH 7, if the two neuraminic acid groups are included) is close to zero under physiological buffer conditions. Considering intrinsic pK values for the ionizable groups of α-thrombin, an overall isoelectric point of 7.8 can be calculated; allowing for the spatial arrangement of the charged groups and their environment (Karshikov et al., 1989) and assuming two neuraminic acid residues located 20 Å apart from the thrombin surface, a higher value of 8.4 is obtained. These calculated isoelectric points are slightly higher than the experimental values of 7.3 and 7.6 obtained by isoelectric focusing (Fenton et al., 1977); however, higher values around 9 have also been reported for human thrombin (Berg et al., 1979; Heuck et al., 1985).

Seventy-nine of the 89 fully charged groups (at pH 7) are at least partially exposed to bulk water (Table 5); they represent 42% of all surface-located residues of the thrombin molecule. This large number of surface charges presumably accounts for the marked tendency of human thrombin to associate at low ionic strength and for its strongly salt-dependent solubility (Landis et al., 1981). The significantly better solubility of the almost equally charged bovine species indicates that some individual surface charges might have a particular impact on the association behavior, however.

Ten charged protein groups of human α-thrombin are completely buried and inaccessible to bulk water (Table 5). All of these buried charges and many of the surface-located charges are arranged in pairs and clusters of oppositely charged groups. Thirty-two salt bridges (of stabilizing energies greater than or equal to $1RT$, i.e.,

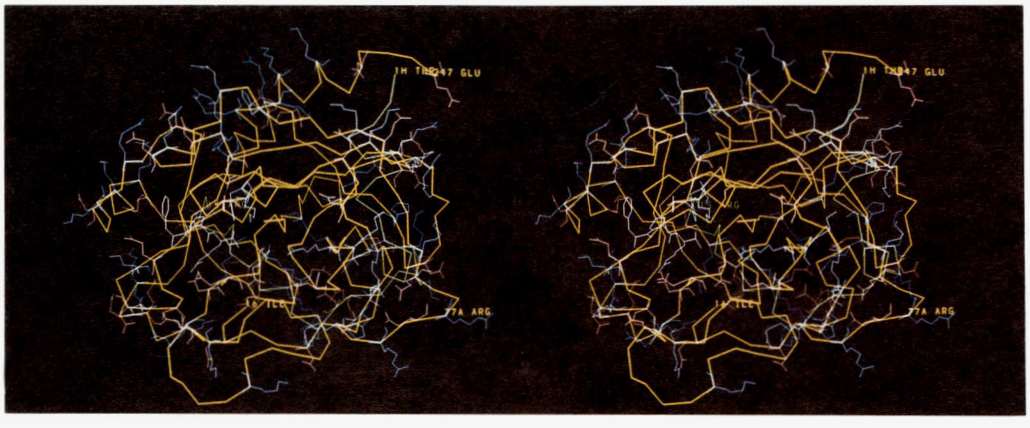

Fig. 7. Main-chain structure of human thrombin (yellow), with all positively (blue) and negatively (red) charged residues added with full structure. Standard view as in Figure 1.

Table 5. *Calculated parameters of all charged groups of human α-thrombin at pH 7, ionic strength 0.125*

Residue	Charge	Electrostatic energy (kcal/mol)	Solvent accessibility of charged group	Residue	Charge	Electrostatic energy (kcal/mol)	Solvent accessibility of charged group
Glu 80	−1.0	−4.6	0.00	Asp 21	−1.0	−0.6	0.65
Lys 202	1.0	−4.4	0.13	Asp 125	−1.0	−0.5	0.35
Asp 14	−1.0	−3.8	0.00	Glu 146	−1.0	−0.5	0.06
NTR[a] 16	1.0	−3.6	0.00	Lys 235	1.0	−0.5	0.36
Arg 187	1.0	−3.6	0.33	Asp 116	−1.0	−0.5	0.71
Arg 137	1.0	−3.2	0.00	Lys 10	1.0	−0.4	0.76
Asp 100	−1.0	−3.2	0.06	Glu 1C	−1.0	−0.3	0.27
Arg 4	1.0	−3.1	0.00	Lys 60F	1.0	−0.3	0.28
Lys 185	1.0	−2.8	0.35	His 119	0.3	−0.2	0.05
Glu 77	−1.0	−2.7	0.12	Lys 145	1.0	−0.2	0.28
Arg 50	1.0	−2.6	0.34	Glu 13	−1.0	−0.1	0.54
Asp 194	−1.0	−2.4	0.00	Arg 173	1.0	−0.1	0.98
Glu 8	−1.0	−2.3	0.00	CTR[b] 247	−1.0	−0.1	0.83
Lys 70	1.0	−2.2	0.00	Asp 60E	−1.0	−0.1	0.79
Asp 1A	−1.0	−2.1	0.36	Asp 170	−1.0	−0.1	0.70
Glu 14C	−1.0	−2.0	0.07	Glu 127	−1.0	−0.1	0.76
Lys 224	1.0	−2.0	0.58	Arg 67	1.0	−0.0	0.21
Lys 135	1.0	−1.9	0.24	His 71	0.1	−0.0	0.26
Arg 221A	1.0	−1.7	0.64	His 91	0.0	0.0	0.18
Glu 14E	−1.0	−1.7	0.24	His 230	0.0	0.0	0.40
Lys 9	1.0	−1.7	0.32	Arg 97	1.0	0.0	0.79
Glu 39	−1.0	−1.7	0.49	Arg 75	1.0	0.0	0.91
Asp 178	−1.0	−1.6	0.68	Glu 186B	−1.0	0.0	0.44
Asp 102	−1.0	−1.6	0.00	Glu 14H	−1.0	0.1	0.73
Glu 61	−1.0	−1.4	0.29	Arg 77A	1.0	0.1	0.73
Lys 14A	1.0	−1.4	0.01	CTR 15	−1.0	0.2	0.81
Glu 86	−1.0	−1.2	0.57	Lys 149E	1.0	0.2	0.51
His 57	0.8	−1.2	0.24	Lys 110	1.0	0.2	0.96
Glu 97A	−1.0	−1.1	0.49	Arg 165	1.0	0.2	0.25
Arg 35	1.0	−1.1	0.47	Asp 14L	−1.0	0.3	0.82
Asp 49	−1.0	−1.0	0.19	Lys 81	1.0	0.3	0.56
Lys 107	1.0	−1.0	0.45	Lys 169	1.0	0.3	0.17
Glu 217	−1.0	−1.0	0.33	Arg 233	1.0	0.3	0.54
Glu 164	−1.0	−0.9	0.41	Lys 36	1.0	0.3	0.36
Lys 87	1.0	−0.9	0.59	Lys 240	1.0	0.3	0.89
Arg 206	1.0	−0.9	0.30	Asp 186A	−1.0	0.3	0.71
Arg 14D	1.0	−0.9	0.87	Glu 192	−1.0	0.4	0.53
NTR 1H	1.0	−0.9	0.36	Lys 109	1.0	0.4	0.53
Glu 247	−1.0	−0.8	0.22	Glu 18	−1.0	0.4	0.89
Lys 186D	−1.0	−0.8	0.51	Arg 101	1.0	0.4	0.19
Asp 63	−1.0	−0.8	0.46	Arg 73	1.0	0.6	0.36
Arg 175	1.0	−0.8	0.35	Arg 126	1.0	0.6	0.72
Asp 243	−1.0	−0.7	0.47	Lys 236	1.0	0.7	0.52
Asp 222	−1.0	−0.7	0.53	Asp 221	−1.0	0.7	0.17
Glu 23	−1.0	−0.6	0.29	Asp 189	−1.0	1.1	0.09
Arg 15	1.0	−0.6	0.66	Arg 93	1.0	1.4	0.72

[a] Amino-terminus.
[b] Carboxy-terminus.

−0.6 kcal/mol) formed between oppositely charged protein groups are given in Table 6. They contribute to the stabilization of both β-barrels and of some of the open turns. Several of the salt bridges are further interconnected giving rise to ionic clusters of up to four protein groups of either charge. The electrostatic energies of the resulting clusters are, of course, slightly less than the arithmetic sum of the contributing ion pairs, due to partial repulsion of identically charged groups.

These electrostatic interactions play particularly important roles in the intrachain stabilization of the A-chain and in the coherence between the A- and the B-chain (see Figs. 9, 10; Table 6). Twenty salt bridges (some of which are further organized in five larger ion pair clusters) are

Table 6. *Ion pairs and clusters stabilizing the tertiary structure of human α-thrombin*

Ion clusters and salt bridges		Electrostatic energy of bridges and clusters (kcal/mol)		Distance[a] (Å)	Solvent accessibility of charged groups involved
Ion pairs stabilizing A-B interactions					
1. Asp 1A–Arg 206		−0.7		6.0/4.2	0.33
2. Glu 8 \ / Lys 202 / Glu 14C \		−1.9 −3.0	(−3.8)	3.8/2.8 2.9/2.7	0.06
3. Asp 14–Arg 137		−3.4		2.8/2.8	0.00
4. Glu 14E / Lys 135 \ Lys 186D		−1.6 −1.1	(−1.6)	3.8/2.8 4.0/3.6	0.33
5. Lys 14A–Glu 23		−1.0		4.5/3.6	0.15
Ion pairs stabilizing A-chain					
1. Asp 1A–Lys 9		−1.8		3.0/2.7	0.34
2. Lys 14A \ / Asp 14 Arg 4 \ Glu 8		−0.8 −2.1 −1.8	(−3.6)	6.0/5.3 3.7/2.9 4.2/2.1	0.0
3. Glu 13–Arg 14D		−0.6		4.4/3.3	0.71
Ion pairs stabilizing B-chain					
1. NTR[b] 16–Asp 194	N[c]	−3.0		3.0/2.7	0.00
2. Arg 35–Glu 39	T	−1.3		3.8/2.7	0.33
3. Arg 50 / Asp 49 \ Glu 247	T C	−1.5 −1.5	(−2.2)	4.0/3.2 4.0/2.7	0.25
4. His 57–Asp 102	A	−1.3		3.9/2.6	0.12
5. Glu 61–Lys 87	S	−1.2		4.1/3.4	0.44
6. Lys 70 \ / Glu 80 Arg 67 \ Glu 77	T	−1.9 −2.5 −2.1	(−4.7)	3.8/3.1 3.3/2.7 3.6/2.5	0.08
7. Glu 86–Lys 107	S	−0.9		4.7/4.5	0.51
8. Asp 100 \ / Arg 175 Glu 97A \ Arg 101	D	−0.9 −0.8 −1.8	(−2.6)	4.1/2.7 5.4/5.2 3.9/3.1	0.27
9. Glu 146–Arg 221A	S	−1.4		3.9/2.8	0.35
10. Glu 164–Lys 185	S	−1.5		3.5/2.7	0.38
11. Arg 233 \ / Asp 178 Arg 165 \	S	−0.6 −0.7	(−1.0)	5.1/3.3 4.9/3.5	0.49
12. Arg 187 / Asp 222 \ Asp 221	S T	−1.7 −1.2	(−2.3)	3.0/2.8 4.5/4.3	0.34
13. Glu 217–Lys 224	S	−1.4		3.7/3.4	0.46

[a] Distances between charged group centers (see Materials and methods)/nearest charged atoms.

[b] Amino-terminus.

[c] Fixation of amino-terminus (N) or carboxy-terminus (C); stabilization of turns (T), β-sheets (S); cross-connection of domains (D) or active-site residues (A).

made by B-chain groups alone. In contrast, the comparatively short A-chain is cross-linked by six salt bridges grouped in three separate clusters; another six salt bridges cross-connect the A- with the B-chain (Fig. 10). Altogether, seven of these salt bridges (two within the A-chain, three between A and B, and two within the B-chain) are completely buried in the protein matrix; in each of these buried ion bridges/clusters, at least one of the charged groups is in direct contact with an adjacent internal water molecule (see Table 9).

Particularly extended ion clusters stretch around Asp 14 (where they confer stabilization to the A-chain and tighten the A–B interaction, see below) and Lys 70 (where they stabilize the structure of loop 70–80 and thus maintain the integrity of the fibrin[ogen] secondary binding exosite; see below). Similar to the other activatable trypsinlike serine proteinases, thrombin possesses a buried salt bridge of large electrostatic interaction energy between the ammonium group of Ile 16 (the amino-terminal residue of the B-chain) and the side-chain carboxyl group of Asp 194 (see below and Kinemage 1); the integrity of this salt bridge has been shown to be important in maintaining the characteristic active conformation (see below and Bode, 1979).

The overall distribution of positively and negatively charged residues in the thrombin molecule and along its molecular surface is not uniform (Fig. 7); rather the positive charges are somewhat more concentrated at two poles, and the negatively charged residues arranged in a ring, leading to a sandwichlike electric-field distribution (Fig. 8). In other words, several charges are not properly balanced by oppositely charged protein neighbors, but cluster in positively and negatively charged surface patches (Table 7). These clusters give rise to quite high electrostatic field strengths (both positive and negative) outside the thrombin surface (Fig. 8). The charged groups of residues located in the center of these patches can exhibit considerable positive electrostatic interaction energies with the electric multipole of the molecule (such as, e.g., Arg 93 or Asp 189 [see Tables 5, 7]) and thus have a destabilizing effect on the molecule; several of the charged residues located more toward the periphery of these patches are, on the other hand, simultaneously engaged in salt bridges and can therefore contribute to the stability of the whole molecule (as, e.g., Glu 39 of the negative patch n1a [Table 7]). This charge clustering is presumably the main reason for the strong electrostatic interaction of thrombin with anionic structures observed even at physiological pH values, where its overall charge almost cancels out.

Three such surface patches (patches p1, p2, n1, and subgroups in Table 7) deserve particular attention. The strongest positive electrostatic field (patch p2) is observed for a groove (on top of the thrombin molecule in Fig. 8) surrounded by the exposed side chains of Arg 126, Lys 236, Lys 240, and Arg 93, and of the electrostatically

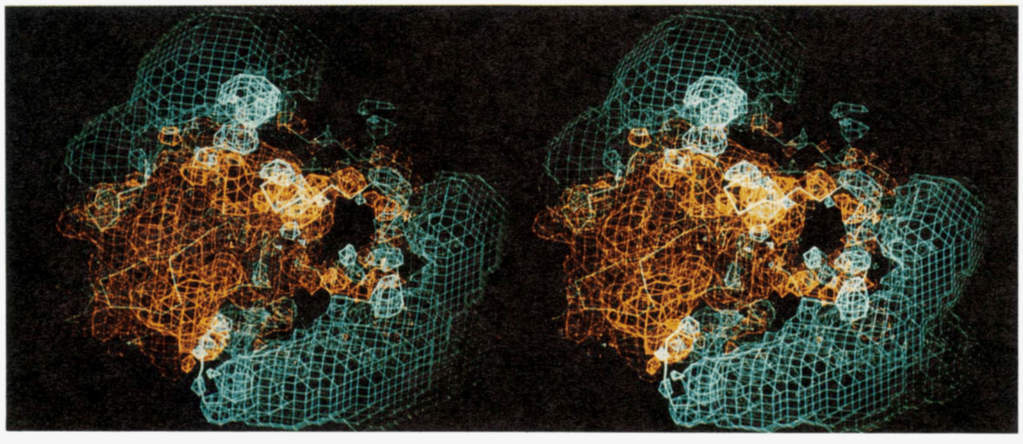

Fig. 8. Main chain structure of human thrombin overlaid with a positive (+3 kcal/mol) and a negative (−5 kcal/mol) electrostatic potential surface. Standard view as in Figure 1.

more compensated residues Lys 235, Arg 101, and Arg 233 in the more distant neighborhood (see Fig. 18). This site presumably represents the heparin-binding site of thrombin and will be discussed further below.

A second surface region of high positive charge density (patch p1) extends along the convex thrombin surface between Lys 149E and Lys 110 (on the right-hand side of the active-site cleft [Fig. 8]). This positively charged patch carries nine basic amino acid side chains (Lys 149E, Arg 67, Arg 73, Arg 75, Arg 77A, Lys 81, Lys 36, Lys 109, and Lys 110) whose charges are only partially compensated by neighboring carboxylate groups. Also this fibrin(ogen) secondary binding exosite will be shown in more detail below.

One larger surface patch of negatively charged residues (Figs. 7, 8) extends around the active site of thrombin (patch n1). It consists of seven aspartic and glutamic acid residues arranged within (Asp 189) and around the specificity pocket: Glu 192 (without protein counter-charge); Asp 102 (active-site residue, hydrogen bonded with His 57); Asp 194 (involved in a relatively strong internal salt bridge with the ammonium group of Ile 16); and Glu 217, Glu 146, and Glu 97A, each involved in salt bridges with positively charged residues in the neighborhood. The many negative charges certainly assist in orienting approaching substrates and affect binding; the electrostatic nature of the catalytic residues located at the edge of this patch is governed more by their adjacent neighbors, however (see below).

Solvent molecules and selected residues

Thrombin possesses four disulfide bridges (see Kinemage 2). The first (Cys 1–Cys 122) links the A- and B-chains, equivalent topologically to that in chymotrypsin(ogen) (Table 8). Disulfide bridge Cys 42–Cys 58 has a similar conformation to that observed in all other vertebrate serine proteinases of known tertiary structure. Cys 168–Cys 182 exhibits the same hand in thrombin as in chymotrypsin but opposite to that in trypsin and kallikrein (Bode et al., 1983); however, this bridge differs in detail from all other related proteinases in spite of a quite similar environment. The fourth disulfide bridge, Cys 191–Cys 220, is again of a similar conformation in thrombin as in other related proteinases.

The 409 (423; see Table 15) localized solvent molecules in the PPACK–thrombin structure account for almost 30% of all solvent (water) molecules in the crystals. Their B-values range from 3 Å^2 to 108 Å^2 (3 Å^2 to 88 Å^2), the average B-value is 50 (45) Å^2. Thirty-four of these solvent molecules (Table 9) are located within the protein domain (and are in direct contact with at most one bulk solvent molecule); they are considered to be an integral part of the thrombin structure. Seventeen of them are found to be topologically equivalent (i.e., with similar protein environment and deviations up to 1.2 Å after optimal superposition of both protein structures) to waters found in benzamidine-free bovine trypsin (DEBA [Bode et al., 1976; see also Table 3 in Bode & Schwager, 1975a]).

Several of these internal solvents are grouped together forming linear water clusters. A particularly large cluster is made by eight adjacent internal solvent molecules (solvent molecules 315, 316, 320, 323, 342, 343, 345, and 414 [Table 9]) encircled by the thrombin insertion loop Tyr 184A–Gly 188 and segment Asp 221–Tyr 225 behind the side chain of Asp 189; Sol 323 mediates the connection to the bulk water (see Fig. 19). Another large internal solvent cluster comprises five water molecules, extending around Sol 360 close to the 70–80 loop. Sol 305 in the back of the specificity pocket (see Fig. 22) is fully separated from bulk solvent molecules in PPACK–thrombin by the Arg 3I side chain; as with the equivalent Sol 416 in trypsin (Bode & Schwager, 1975a), it plays an impor-

Table 7. *Charged groups of human α-thrombin that form charge patches*

Charge patch	Residue	Interaction energy (kcal/mol)	Nearest charged neighbor (distance and interaction energy) Å/residue/(kcal/mol)
Positive patches			
p1	Lys 36	0.3	(6.7/Lys 109/0.4)
	Arg 73	0.6	(8.2/Lys 149E/0.3)
	Arg 75	0.0	(6.6/Glu 77/−0.5)
	Arg 77A	0.1	(>10./)
	Lys 81	0.3	(9.9/Lys 110/0.1)
	Lys 109	0.4	(6.7/Lys 36/0.4)
	Lys 110	0.2	(9.4/Lys 109/0.1)
	Lys 145	−0.2	(>10./ /)
	Lys 149E	0.2	(6.6/Asp 21/−0.3)
p2a	Arg 93	1.4	(3.6/Arg 101/1.3)
	Arg 101	0.4	(3.6/Arg 93/1.3)
	Arg 126	0.6	(4.3/Lys 236/0.4)
	Arg 233	0.3	(4.9/Asp 178/−0.7)
	Lys 235	−0.5	(6.8/Asp 125/−0.5)
	Lys 236	0.7	(4.3/Arg 126/0.4)
	Lys 240	0.3	(9.7/Lys 236/0.1)
p2b	Arg 165	0.2	(4.5/Lys 169/0.8)
	Lys 169	0.3	(4.5/Arg 165/0.8)
p2c	Arg 97	0.0	(>10./ /)
	Arg 173	1.0	(8.6/Glu 217/−0.2)
	Arg 195	1.0	(4.1/Glu 97A/−0.9)
Negative patches			
n1a	Glu 39	−1.7	(3.8/Arg 35/−1.3)
	Asp 60E	−0.1	(>10./ /)
	Glu 61	−1.4	(4.1/Lys 87/−1.2)
	Asp 63	−0.8	(7.4/Lys 36/−0.4)
	Glu 97A	−1.1	(4.1/Arg 175/−0.9)
	Glu 146	−0.5	(3.9/Arg 221A/−1.4)
	Asp 189	1.1	(6.6/Asp 221/0.7)
	Glu 192	0.4	(7.7/Glu 146/0.4)
	Glu 217	−1.0	(3.7/Lys 224/−1.4)
	Asp 221	0.7	(4.5/Arg 187/−1.2)
n1b	Glu 18	0.4	(6.0/Asp 221/0.6)
	Glu 164	−0.9	(3.5/Lys 185/−1.5)
	Asp 170	−0.1	(8.4/Lys 169/−0.2)
	Asp 186A	0.3	(6.4/Glu 186B/0.4)
	Glu 186B	0.0	(5.2/Lys 185/−0.8)
	Asp 222	−0.7	(3.0/Arg 187/−1.7)
n2	Asp 1A	−2.1	(3.0/Lys 9/−1.8)
	Glu 1C	−0.3	(7.6/Arg 206/−0.5)
	Asp 49	−1.0	(4.0/Arg 50/−1.5)
	Glu 247	−0.8	(4.0/Arg 50/−1.5)
	CTR[a] 247	−0.1	(5.8/Glu 247/0.5)
n3	Glu 14H	0.1	(7.4/Glu 14E/0.4)
	Asp 14L	0.3	(6.8/CTR 15/0.2)
	CTR 15	0.2	(6.8/Asp 14L/0.2)
	Glu 127	−0.1	(7.3/Asp 125/0.2)

[a] Carboxy-terminus.

tant role as an anchor for the distal positively charged groups of substrate P1 residues.

The aromatic residues are relatively evenly distributed over the thrombin molecule. Edge-on-plane as well as planar stacking of adjacent aromatic rings can be observed (e.g., Tyr 60A–Trp 60D and Phe 225–Trp 215). The concentration of four relatively exposed indole side chains (Trp 60D, Trp 148, Trp 96, and Trp 215) around the substrate-binding site is particularly remarkable (see below).

Table 8. *Dihedral angles of the disulfide bridges in thrombin*

Disulfide bridge	κ1	κ2	κ3	κ4	κ5	Hand
Cys 1–122	68	77	97	−63	179	Right
Cys 42–58	−82	−163	−95	−78	−77	Left
Cys 168–182	−62	−61	−90	−77	90	Left
Cys 191–220	−157	46	88	−170	−55	Right

The thrombin A-chain and its interaction with the B-chain

The thrombin A-chain with its multiple-turn/helical conformation (see Fig. 3) is fixed to the B-chain surface opposite to the substrate-binding cleft (Fig. 2; Kinemage 3). It has an overall boomeranglike shape and nestles in such a way toward the B-chain that it merges to the common molecular surface without significant breaks (with the exception of the extending carboxy-terminal segment and the exposed elbow around Lys 10–Ser 11 and Glu 13–Arg 14D [Fig. 9]). The buried surface between the chains amounts to 1,200 Å^2; this number can be compared to the intermolecular surface of 1,725 Å^2, buried between α-thrombin and hirudin in their complex (Rydel et al., 1991).

Besides the covalent disulfide bridge connection between Cys 1 and Cys 122 the interaction of the A- with the B-chain is mediated through charged side chains. Only three inter-main-chain hydrogen bonds exist between A and B (see Fig. 3), whereas another 20 hydrogen bonds are formed utilizing at least one side chain (Fig. 10). Six of the latter represent simultaneously (largely) buried salt bridges (see Table 6), and a charged side-chain group participates in a further 10 (such hydrogen bonds involving charged groups are particularly strong).

Stabilization within the A-chain occurs preferentially through polar and salt-bridge interactions (in particular in its central part [see Figs. 9, 10]). Fourteen of the intra-A-chain hydrogen bonds are made between main-chain groups (Fig. 3). Nine oppositely charged side-chain groups are involved in four of the six residual hydrogen bonds and form part of the extended salt-bridge clusters (Table 6); three of them contribute simultaneously to the interchain salt bridges mentioned above (Fig. 10).

The amino-terminal segment of the A-chain up to Glu 1C is only defined by relatively weak density (see Fig. 28) in the PPACK–α-thrombin crystals but is placed unequivocally. This segment runs along a ridge of the B-

Table 9. *Internal water molecules and their hydrogen-bond partners as found in thrombin PPACK*[a]

PPACK-human α-thrombin					Bovine trypsin		PPACK-human α-thrombin					Bovine trypsin	
	Hydrogen-bond partners							Hydrogen-bond partners					
Solvent number	Atom		Distance (Å)	B (Å^2)	Solvent number	Deviation of equivalent molecules (Å)	Solvent number	Atom		Distance (Å)	B (Å)	Solvent number	Deviation of equivalent molecules (Å)
304	Asn 143	Nδ2	(2.9)	12	–	–	325	Met 26	O	(2.7)	15		
	Lys 145	O	(3.2)		–	–		Arg 137	NH	(3.0)			
	Thr 147	N	(3.2)		–	–		Sol 454	OH	(2.7)			
305	Trp 215	O	(3.4)	4	416	0.9	327	Pro 5	N	(3.5)	31	–	–
	Phe 227	O	(3.3)					Leu 6	N	(3.1)			
	Arg 31	NH1	(3.0)					Gly 25	O	(2.6)			
	Tyr 228	OH	(3.7)					Asp 116	O	(2.8)			
306	Met 210	O	(3.0)	14	722	0.7	332	Arg 101	N	(2.9)	28	703	0.3
	His 230	Nδ1	(3.2)					Asp 102	N	(3.5)			
	Phe 232	N	(2.9)					Asn 179	O	(2.9)			
308	Ser 45	Oγ	(3.1)	13	406	0.7		Sol 368	OH	(2.7)			
	Leu 53	O	(2.8)				334	Ser 171	Oγ	(3.0)	15	–	–
	Gly 196	O	(2.6)					Tyr 225	N	(2.9)			
	Gln 209	Nε2	(2.9)					Sol 411	OH	(2.9)			
310	Glu 217	O	(2.7)	8	415	0.3	335	Trp 148	Nε1	(2.9)	23	–	–
	Gly 219	O	(3.5)					Glu 192	Oε1	(3.1)			
	Cys 220	N	(3.4)					Cys 220	N	(3.4)			
	Arg 221	N	(3.0)					Cys 220	Sγ	(3.2)			
311	Trp 141	N	(3.1)	19	410	0.4	337	Glu 217	N	(2.9)	22	–	–
	Gly 193	O	(2.8)					Tyr 225	O	(2.6)			
	Sol 314	OH	(3.2)					Sol 310	OH	(3.4)			
314	Glu 30	Nε2	(3.2)	10	701	0.1		Sol 411	OH	(3.3)			
	Thr 139	O	(2.7)				342	Arg 221	O	(2.7)	10	–	–
	Asp 194	O	(3.0)					Lys 224	O	(2.8)			
	Sol 311	OH	(3.2)					Sol 316	OH	(2.8)			
315	Ala 183	O	(3.0)	16	704	0.3		Sol 343	OH	(2.4)			
	Asp 189	Oδ2	(2.8)					Sol 414	OH	(3.3)			
	Gly 226	N	(3.5)				343	Tyr 184	O	(2.9)	10	–	–
	Tyr 228	OH	(2.8)					Lys 224	O	(2.6)			
316	Asp 221	Oδ1	(2.7)	11	–			Sol 320	OH	(3.1)			
	Sol 342	OH	(2.5)					Sol 342	OH	(2.4)			
	Sol 320	OH	(2.8)					Sol 345	OH	(2.6)			
318	Glu 30	O	(3.0)	17	708	0.2	345	Gly 188	O	(2.8)	27	705	0.5
	Arg 67	O	(2.7)					Sol 315	OH	(3.3)			
	Gly 69	N	(3.3)					Sol 320	OH	(3.1)			
	Lys 70	N	(3.2)					Sol 343	OH	(2.6)			
319	Ile 16	N	(2.7)	3	430	0.3	359	Tyr 117	O	(3.0)	39	–	–
	Gly 140	O	(3.1)					Sol 358	OH	(3.3)			
	Gly 142	N	(2.9)					Sol 360	OH	(2.7)			
	Gly 142	O	(2.7)					Sol 457	OH	(3.3)			
320	Tyr 184	N	(2.9)	13	–		360	Ser 27	O	(3.4)	15	709	1.2
	Gly 188	O	(3.0)					Lys 70	O	(2.8)			
	Sol 316	OH	(2.8)					Sol 322	OH	(3.0)			
	Sol 343	OH	(3.1)					Sol 359	OH	(2.7)			
	Sol 345	OH	(3.1)					Sol 457	OH	(3.0)			
321	Arg 137	O	(3.2)	21	721	0.4	368	Leu 99	O	(2.8)	16	408	0.4
	Thr 139	Oγ	(2.8)					Ser 214	Oγ	(2.9)			
	Pro 198	O	(3.0)					Sol 322	OH	(2.7)			
	Sol 360	OH	(3.0)				386	Val 241	O	(2.6)	21	–	–
323	Lys 185	O	(2.8)	28	–	–		Ile 242	O	(2.8)			
	Lys 186D	O	(3.0)					Gly 244	N	(3.2)			
	Sol 324	OH	(2.9)					Phe 245	N	(3.2)			

(*continued*)

Table 9. *Continued*

PPACK-human α-thrombin					Bovine trypsin	
	Hydrogen-bond partners					
Solvent number	Atom		Distance (Å)	B (Å²)	Solvent number	Deviation of equivalent molecules (Å)
410	Arg 165	Nε	(2.7)	25	–	–
	Arg 165	NH2	(3.2)			
	Thr 177	O	(3.0)			
	Met 180	O	(3.1)			
411	Glu 217	Oε2	(3.0)	25	–	–
	Sol 334	OH	(2.9)			
	Sol 337	OH	(3.3)			
414	Tyr 184	O	(2.7)	21	–	–
	Gly 223	N	(2.7)			
	Sol 342	OH	(3.3)			
454	Ser 27	Oγ	(3.3)	25	–	–
	Trp 29	Nε1	(3.0)			
	Sol 321	OH	(2.9)			
	Sol 325	OH	(2.7)			
457	Ser 27	N	(2.9)	35	717	0.7
	Ser 27	O	(3.3)			
	Sol 359	OH	(3.3)			
	Sol 360	OH	(3.0)			
626	Trp 51	Nε1	(3.0)	44	713	0.8
	Lys 107	Nζ	(3.4)			

[a] Topologically equivalent solvent molecules of bovine trypsin (benzamidine-free, DEBA [Bode et al., 1976]) are added together with the r.m.s. deviation obtained after optimal superposition of both enzyme models (for deviations up to 1.2 Å).

chain with only a few van der Waals contacts between both chains. The position of the amino-terminal Thr 1H is almost in the center of the carboxy-terminal (nonhelical) half-turn of the B-chain (see Fig. 9). It is close to the negatively charged side chains of Asp 243 and Glu 247 but not within direct hydrogen bond distance (see Table 5).

The equivalent A-chain segment seems to be similarly arranged in bovine thrombin; in the crystals of human thrombin bound to hirudin (Rydel et al., 1991) it is definitely in a quite different position and conformation, presumably due to intermolecular contacts. The A-chain amino-terminus therefore possesses a high degree of flexibility and can adapt to different crystal environments but seems inherently to prefer the location found in PPACK-thrombin. The preceding segment, cleaved off autocatalytically upon activation in human prothrombin, might extend away from the molecular surface thus facilitating enzymatic attack; in bovine thrombin the equivalent 13-residue segment, which is not cleaved off in the course of activation, is completely undefined by electron density, indicating its complete disorder.

The central, more rigid part of the A-chain (between Asp 1A and Tyr 14J) runs in a shallow curved groove of the B-chain lined by B-chain segments Ser 20–Met 26 and Lys 202–Trp 207 (Fig. 9). The amino-terminal half of this central segment, organized mainly in tightly packed turns, is stabilized considerably by several intrachain and interchain hydrogen bonds (Fig. 3) and by salt-bridge clusters (Table 6). Asp 14 and Lys 202, with their side chains pointing toward the molecular center and toward the peripheral A-chain, respectively, are at the center of the charge network (Fig. 10). The carboxylate group of Asp 14 opposes the guanidyl group of (B-chain residue) Arg 137 almost symmetrically; furthermore, it is involved in favorable hydrogen bond/salt bridge interactions with the guanidyl group of (A-chain residue) Arg 4. Glu 8 makes hydrogen-bond/electrostatic interactions with this Arg 4, but also with Lys 202, which further interacts with the carboxylate group of Glu 14C.

Most of these salt bridges are more or less completely buried (Table 6). Electrostatic interactions therefore contribute significantly to the intrachain stability as well as to the A–B interaction; almost 90% of the total electrostatic energy of the interaction between A and B is calculated to be due to these salt clusters. Six out of 7

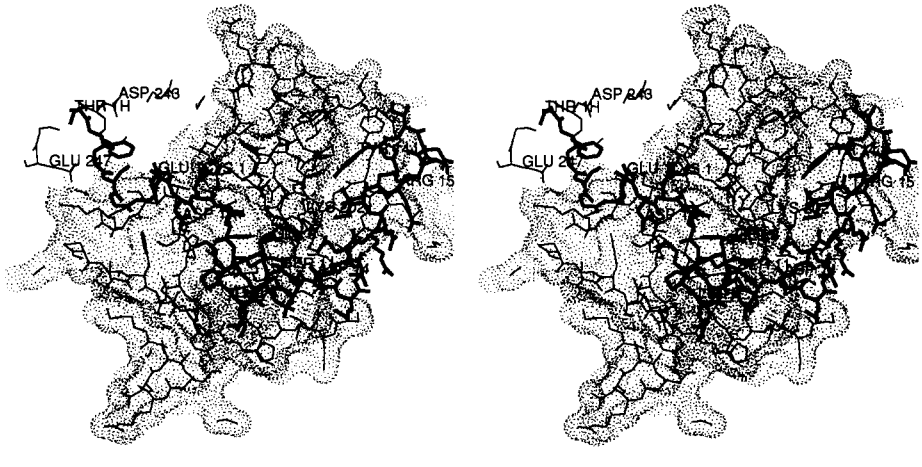

Fig. 9. Back view of the human thrombin B-chain (thin connections) and its Connolly surface, displayed together with the thrombin A-chain (thick connections). This figure is obtained from Figure 1 after an approximately 180° rotation about a vertical axis. The A-chain runs from left (close to the B-chain carboxy-terminus Glu 247) to right along a boomeranglike shallow groove of the B-chain.

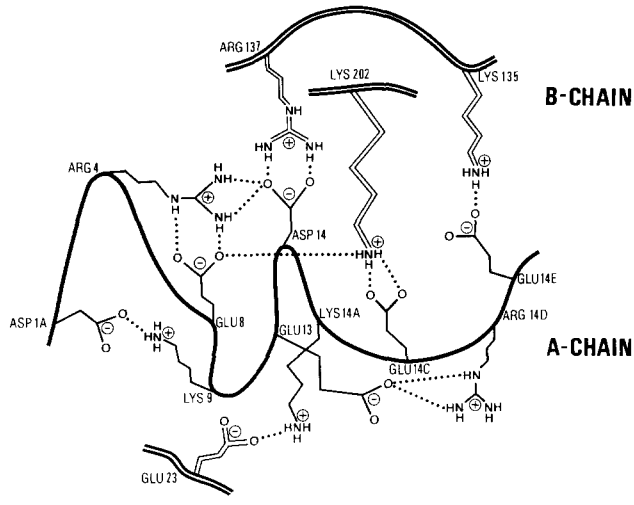

Fig. 10. Schematic drawing of the ionic interactions (as far as involved in hydrogen bonding) within the thrombin A-chain and between the thrombin A-chain (bold connections) and the B-chain (double connections). The arrangement is similar to that in Figure 9, i.e., corresponds to a back view of the thrombin molecule.

A-chain residues involved in intra- and interchain salt bridges are shared by human and bovine thrombin; only Glu 13, which makes a weak surface-located salt bridge in human thrombin, is replaced by a glutamine in the bovine species. In spite of these impressive electrostatic interactions, however, the A-chain is apparently neither of great importance for B-chain folding nor for the (thermodynamic) stability of the B-chain globule (Hageman et al., 1975).

The carboxy-terminal A-chain segment between Thr 14B and Tyr 14J is organized in an amphiphilic α-helix of $1\frac{1}{2}$ turns (Figs. 3, 9). The side chains of Leu 14G and Tyr 14J are involved in hydrophobic contacts between the A- and the B-chains. The phenolic side chain of Tyr 14J exhibits enhanced flexibility. The segment from Ile 14K to the carboxy-terminal Arg 15, which projects away from the main molecular body, is only weakly defined by electron density. The carboxy-terminus at Arg 15 is positioned close to the guanidyl group of Arg 165 of a symmetry-related molecule. Segment Ile 14K–Arg 15 is arranged quite differently in bovine thrombin as well as in the human thrombin–hirudin complex; it therefore appears to be inherently flexible.

A-chain segment Cys 1–Leu 6 is topologically equivalent to the first six amino acid residues of the activation peptide of chymotrypsin(ogen) and connected in a similar manner via disulfide bridge Cys 1–Cys 122 to the thrombin B-chain. The more carboxy-terminal part of the thrombin A-chain is of a different conformation and considerably longer than the corresponding pro-part of chymotrypsinogen. The interaction surface of the thrombin B-chain has a relatively similar contour to that of chymotrypsinogen; in thrombin, however, most of the interacting side chains are longer and more polar (see Fig. 11).

In human α-thrombin, Arg 15 is about 20 Å away from the equivalent position in chymotrypsinogen (Fig. 11). Assuming an A-chain conformation identical to α-thrombin up to Tyr 14J, Arg 15 of a prothrombin molecule would not be able to attain an equivalent position, even after a complete extension of the connecting polypeptide segment. Presumably, therefore, thrombin segment Ile 16–Ser 20 occupies a different position in prothrombin.

As mentioned above, the B-chain alone appears to be thermodynamically stable and to retain clotting activity (Hageman et al., 1975). Thus, the A-chain would not seem to be a structural element required for the activity of thrombin; its integration into the B-chain might rather be necessary for correct activation cleavage and for other functions of prothrombin. As it is positioned opposite to the active site, interaction of the A-chain with larger bound substrates and inhibitors is not possible. It is, however, conceivable that amino-terminally linked pro-parts,

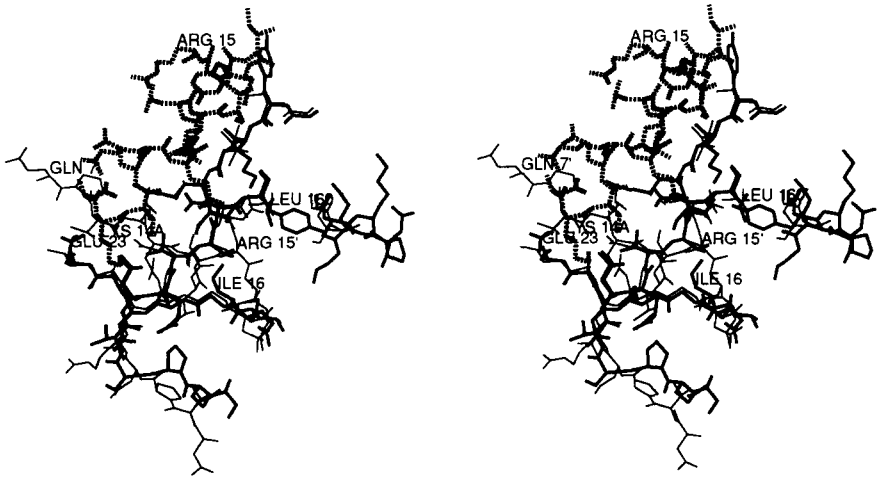

Fig. 11. Parts of the thrombin A-chain (dashed connections) and B-chain (thick connections) optimally superimposed with equivalent parts of bovine chymotrypsinogen (thin connections). Polypeptide segment 10–20 of chymotrypsinogen deviates considerably from the equivalent parts of both thrombin chains; the carboxy-terminal residue Arg 15 of the thrombin A-chain is quite distant from the equivalent Arg 15′ of the (nonactivated!) chymotrypsinogen. This back view is similar, but not identical, to that used in Figure 9.

such as fragment 1.2 of meizothrombin, could fold back and project into the substrate-binding region, thus impairing activity toward macromolecular substrates, as recently shown for fibrinogen or Factor V (Doyle & Mann, 1990).

The thrombin B-chain

The tertiary structure of the thrombin B-chain is homologous to the reactive domains of the other trypsinlike serine proteinases of known spatial structure (see Figs. 2, 4, 5 and Bode et al., 1989b). A remarkable feature of the thrombin B-chain globule is, however, a series of more elongated and exposed loops, in particular around the active-site cleft. The unique specificity of thrombin for macromolecular ligands involves interactions at the surface of the molecule. The surface segments of the thrombin B-chain exhibit new features compared with chymotrypsinogen (as can be inferred from Table 3, Figs. 4, 5 and as has been presented in more detail by Bode et al., 1989b). Among them are (1) segment Arg 173–Ile 174; (2) loop segment 95–100; (3) the large insertion loop Leu 59–Asn 62 (the 60-insertion loop or Trp 60D loop), and (4) loop segment 34–42, which together form the "north rim" of the active-site canyon (Bode et al., 1989b); (5) the large insertion loop 145–150, which together with the trypsinlike segment Gly 216–Cys 220 forms the corresponding "south rim"; (6) loop segment Tyr 184A–Gly 188, and (7) the trypsinlike loop segment Cys 220–Tyr 225, which together make up the thrombin structure behind the specificity pocket; and (8) loop segment Ser 203–Arg 206, which is particularly involved in the interaction with the A-chain (see above).

Most of these surface loops project out of the molecular surface. However, with the exception of loop segment 145–150 (see above and below), all of these loops are remarkably rigid. Proline and tryptophan residues seem to contribute to their rigidity.

The positions of some polypeptide segments of the α-thrombin B-chain particularly susceptible to proteolytic cleavage are emphasized in Figure 12 and Kinemage 2. Some of these cleavage sites (in particular those generated autocatalytically or through trypsin action) may only exist under laboratory conditions and might play no important role in vivo; they are, nevertheless, extremely valuable for investigating the importance of distinct thrombin sites for various thrombin functions. Most of these sites are located at exposed surface loops. One of them (Lys 149–Gly 150, see below) is not readily accessible to an attacking proteolytic enzyme; another one (Arg 67–Ile 68) is even completely buried in the interior of the thrombin molecule and only accessible after prior cleavage in the 70–80 loop (see below). None of the preferred proteolytic cleavage sites shown in Figure 12 exhibit the canonical conformation found in substratelike binding small protein inhibitors (Huber & Bode, 1978; Laskowski & Kato, 1980) (see Table 10). Most of these sites are, however, flexible enough to allow conformational adaptation to the binding sites of the proteolytic enzymes. The rigid, noncomplementary activation cleavage site of chymotrypsinogen (Wang et al., 1985), however, illustrates nicely that the susceptibility of peptide bonds is governed by other (hitherto unknown) properties.

Thrombin sites of particular importance

Ile 16–Asp 194 and its role in activation

Similar to the other activatable trypsinlike serine proteinases, the amino-terminal residue Ile 16 (and part of

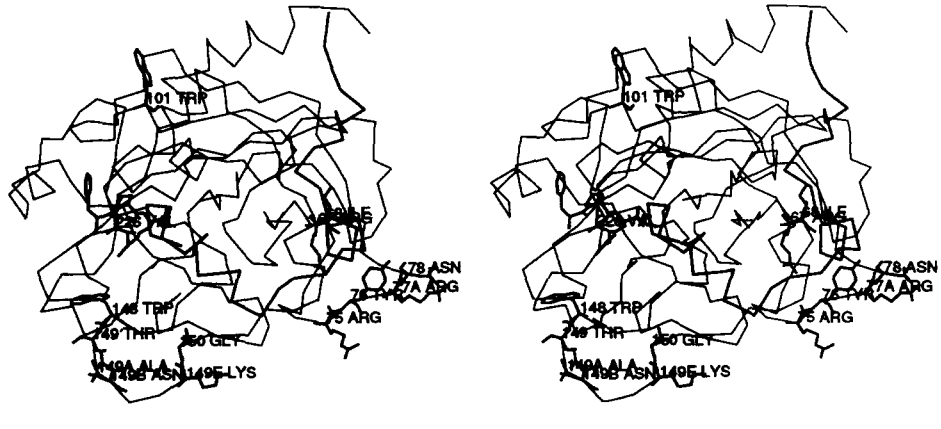

Fig. 12. α-carbon chain of the A-chain (thick connections) and the B-chain (thin connections) of human α-thrombin displayed together with the full amino acid residues placed at well-characterized proteolytic cleavage or replacement sites in naturally occurring thrombin variants. The bound PPACK is overlaid. Cleavages at Arg 77A–Asn 78 and Arg 67–Ile 68 give rise to β-thrombin; cleavages at Trp 148–Thr 149, Ala 149A–Thr 149B, and Lys 149E–Gly 150 yield ζ-, ϵ-, and γ-thrombin, respectively. In the genetic thrombin variants Quick I, Quick II, and Tokushima, Arg 67, Gly 226, and Arg 101 are replaced by cysteine, valine, and tryptophan, respectively. Standard view as in Figure 1.

Table 10. *Main-chain conformational angles of some α-thrombin cleavage sites compared with the proteinase-binding loops of some protein inhibitors around the scissile peptide bond*

Scissile peptide bond	P2	P1	P1′	P2′
β_t: Arg 77A–Asn 78	−98/95	−48/−55	−115/29	−114/−52
β': Arg 75–Tyr 76	−124/−30	−85/151	−67/118	−98/95
β: Arg 67/Ile 68	−130/124	−111/118	−115/143	79/8
γ_t: Lys 149E–Gly 150	110/164	−75/149	−92/−22	−127/137
ζ: Trp 148–Thr 149	−101/−6	−73/−8	−64/128	−163/154
ϵ: Ala 149A–Asn 149B	−137/150	−64/128	−163/154	−100/153
BPTI: Arg 15I–Ala 16I	−82/160	−120/34	−88/179	−132/84
Eglin c/subtilisin: Leu 45–Asp 46	−61/143	−115/43	−96/169	−117/110
OMTKY3: Leu 18I–Glu 19I	−76/155	−99/29	−82/156	−104/106
PPACK	−74/139	−87/		

Val 17) of the thrombin B-chain is buried within the protein moiety (see Kinemage 1). The internal salt bridge made between the α-ammonium group of Ile 16 and the side-chain carboxylate group of Asp 194 is of considerable strength (3.0 kcal/mol [see Table 6]). According to our electrostatic calculations, the pK of this ammonium group is shifted to about 10.5, due mainly to this strong interaction. The deprotonation of this ammonium group and the concomitant disruption of the salt bridge is generally believed to result in a transformation of the active enzyme structure toward a zymogen-like structure (for a more detailed discussion of the equivalent case in trypsin, see Bode [1979]). This transformation is probably accompanied by a decrease in activity with increasing pH values in the alkaline range. Indeed, the thrombin hydrolytic activity toward typical synthetic *p*-nitroanilide substrate seems to be governed by an ionizable group of a pK value of almost 10 (see Lottenberg et al., 1983; De Cristofaro & DiCera, 1990).

In prothrombin the intact Arg 15–Ile 16 segment must be situated on the molecular surface (see Fig. 11). It would be reasonable to assume that segment 191–194 and the adjacent residues in prothrombin might have a similar conformation as the equivalent residues in chymotrypsinogen (Freer et al., 1970; Wang et al., 1985); i.e., that Asp 194 might project into the active-site cleft and loop segment 191–194 folded inward with residue 192 occupying the specificity pocket. However, the thrombin residue topologically equivalent to His 40 of trypsinogen and chymotrypsinogen is Leu 40; thus, Asp 194 of prothrombin could not, as in the chymotrypsinogen structure, accept a (probably stabilizing) hydrogen bond from the side chain of residue 40. In addition, according to model-building experiments the polar side chain of Glu 192 of prothrombin would probably not find a favorable environment in the protein interior of prothrombin (in contrast to the equivalent methionine side chain of chymotrypsinogen). Therefore, both substitutions (at positions 40 and 192) seem to favor destabilization of the thrombin zymogen state relative to the active state. As mentioned earlier (Bode, 1979), such an equilibrium shift toward the active-state-like conformation could facilitate formation of proteolytically active prothrombin–protein complexes such as the prothrombin–staphylocoagulase complex (see Kawabata et al., 1985).

The insertion loop Leu 59–Asn 62 (60 insertion loop)

The polypeptide segment Cys 58–Leu 64 of thrombin is nine residues longer than the corresponding segment of chymotrypsin (see Table 3; Fig. 4) and folded in a characteristic hairpin loop, which projects (in contrast to the digestive enzymes) considerably out of the molecular surface (Figs. 4, 5, 13A). This large insertion loop plays an important role in the selection of susceptible (macromolecular) substrates and inhibitors, i.e., for recognition and specificity of thrombin (Bode et al., 1989b). Segment Leu 59–Phe 60H is folded in an approximate 3_{10}-helical turn, which is terminated by Glu 60E in a left-handed helical conformation (Table 2). The phenolic side chain of Tyr 60A points toward the S2 subsite, covering it, and packing tightly against Pro 2I of the bound PPACK molecule (see below). The indole moiety of Trp 60D is almost fully exposed to bulk solvent molecules. This part of the exposed loop seems to be particularly stabilized by the two intervening proline residues, which, in addition, pack against Trp 96 of an adjacent loop (see Fig. 13A and Kinemage 4). In the PPACK–thrombin crystals the whole loop (including all side chains) is well defined by electron density (Fig. 13B) and also makes intermolecular contacts with residues of the B-chain carboxy-terminus of a symmetry-related thrombin molecule. In spite of very different crystal contacts in the thrombin–hirudin (Rydel et al., 1991) and in the bovine thrombin structure, this loop is of similar structure, i.e., its internal conformation seems to be stable.

The well-defined (carbohydrate carrying) asparagine side chain of the subsequent Asn 60G extends from the loop surface away from the active-site cleft (Fig. 13A). In the PPACK–thrombin crystals there is free intermo-

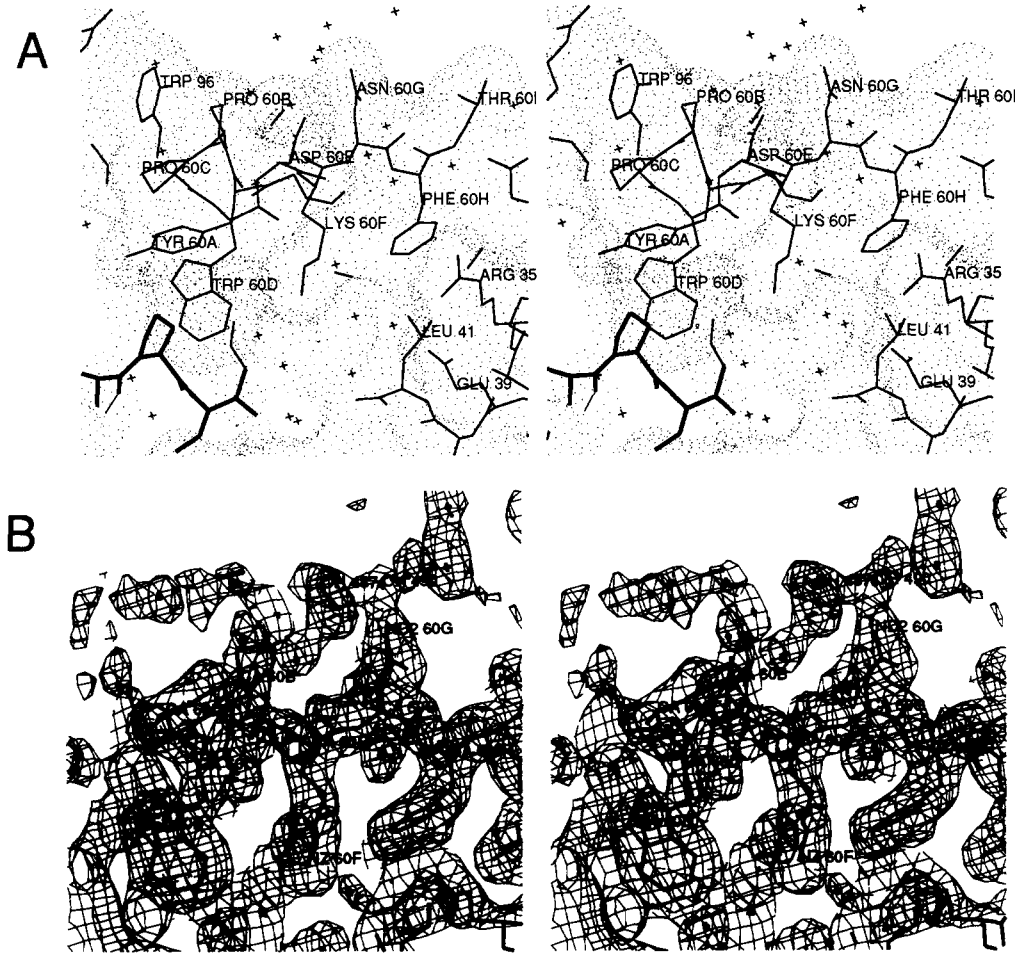

Fig. 13. A: The 60 insertion loop and the glycosylation site of human α-thrombin (thin connections) displayed together with part of PPACK (thick connections) and the thrombin Connolly surface. Loop segment Tyr 60A–Asp 60E projects particularly out of the molecule; it probably forms part of the chemotactic domain of thrombin. Asn 60G is pointing away from the active-site cleft. Standard view as in Figure 1. **B**: Section of the final electron density around the 60 insertion loop superimposed with the equivalent parts of human α-thrombin and some localized solvent molecules (crosses). The electron density representing the Asn 60G side chain (upper right corner) extends toward the top; it is, however, not appropriate to accommodate a distinct Asn 60G *N*-linked sugar chain; instead, this density has been interpreted as two fixed solvent molecules. Contour surface at 0.8σ calculated for the whole map. Standard view as in Figure 1.

lecular space but no continuous electron density that would indicate the carbohydrate linked to it (Fig. 13B); the isolated density blobs visible beyond the carboxamide group have therefore been interpreted as solvent molecules. The bound, presumably disordered, oligosaccharide (consisting of 12 sugar units in a branched chain [Nilsson et al., 1983]) must extend away from the active-site cleft (Fig. 13A). Thus, it should not interfere with macromolecular substrates that bind mainly along the cleft; this is in agreement with the lack of any effect on enzymatic (clotting and esterase) and cellular activities of thrombin on its removal or destruction (Horne & Gralnick, 1983). Concanavalin A, in binding to the α-D-mannose of this sugar chain, might on the other hand come into conflict with an approaching fibrinogen molecule, in agreement with experimental results (Skaug & Christenson, 1971; Karpatkin & Karpatkin, 1974; Hageman et al., 1975).

The compact and relatively rigid segment Tyr 60A–Trp 60D is part of a tetradecapeptide with the sequence Tyr 60A–Leu 65 of known extracellular matrix-binding and chemotactic/growth factor activities (Bar-Shavit et al., 1984, 1986, 1989). The intact catalytic residues are not a prerequisite for chemotactic activity toward monocytes and neutrophils (Bizios et al., 1986). It is conceivable that such activities could be attributed to this exposed Tyr 60A–Trp 60D loop of α-thrombin for its rather unique ridgelike architecture. In thrombin–hirudin this projecting thrombin loop is smoothed out by the hirudin Pro 46HI–Glu 49HI segment packing alongside of it (Rydel et al., 1991); thus, partial masking upon hirudin binding could account for the observed abolition of chemotactic

activity (Bizios et al., 1986; Prescott et al., 1990). Extracellular matrix-bound thrombin failed to form a complex with antithrombin III, but still exhibits an open proteolytic site (Bar-Shavit et al., 1989). This is in agreement with recent model-building studies, which show that the core of antithrombin III should pack against this projecting loop thus burying it and removing it from contacts with receptors (unpubl.).

According to a search in 7,068 sequences, this YPPW sequence seems (with the exception of a presumed polymerase subunit from a vesicular stomatitis virus) to be confined to thrombin. Preliminary molecular dynamics simulations with an isolated and *N*-methylated Tyr-Pro-Pro-Trp-Asn-amide peptide indicate that the distinct conformation observed in thrombin might also be a low free-energy conformation in the isolated peptide; recent NMR experiments show that the equivalent hexapeptide does, however, not possess a single preferred conformation but interchanges between at least two conformational states (unpubl.).

Surface structure around segment 67–80 (the fibrin(ogen) secondary binding exosite)

The loop Lys 70–Glu 80 (Fig. 14A; Kinemage 5) of thrombin is topologically similar to the so-called calcium-binding loop of bovine trypsin (Bode & Schwager, 1975a,b) (Fig. 15A) and of the other digestive serine proteinases (see Meyer et al., 1988). The central residue of this loop is Lys 70. Its distal ammonium group occupies a site equivalent to the calcium ion in the pancreatic enzymes (Fig. 15A). In the thrombin structure, however, of the six oxygen ligands of calcium found in bovine trypsin only three are connected to Lys 70Nζ through short hydrogen bonds, in an almost tetrahedral geometry (Table 11). It is noteworthy, that the distal ammonium group of Lys 70, which is well below the molecular surface in

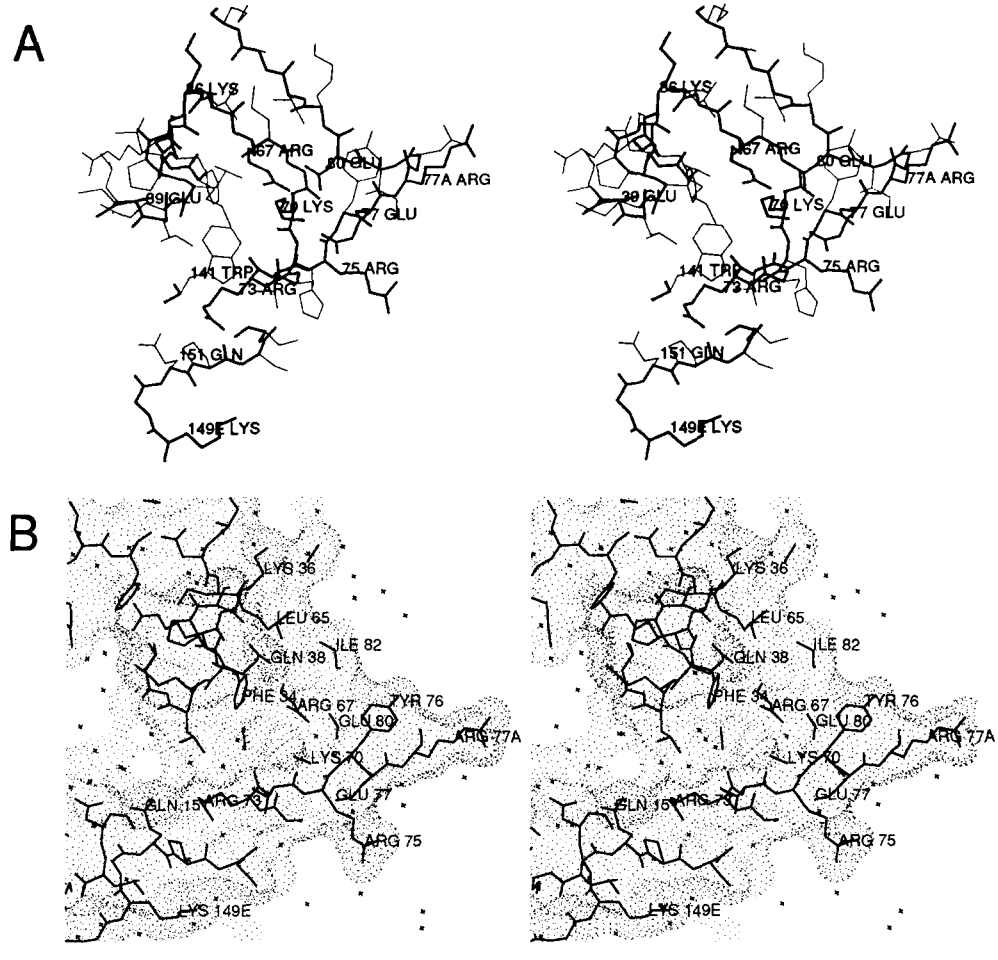

Fig. 14. The fibrinogen secondary binding exosite of human α-thrombin. The view is similar to the standard view of Figure 1. **A**: Stick drawing of loop 67–80 and its environment. Lys 70 is located at the center of an extended buried salt bridge network made with Glu 77, Glu 80, and Arg 67. Cleavage of peptide bond Arg 77A–Asn 78 leads to formation of noncoagulant β_1-thrombin, in which this loop is presumably unfolded. **B**: Surface residues of the exosite displayed together with the Connolly surface.

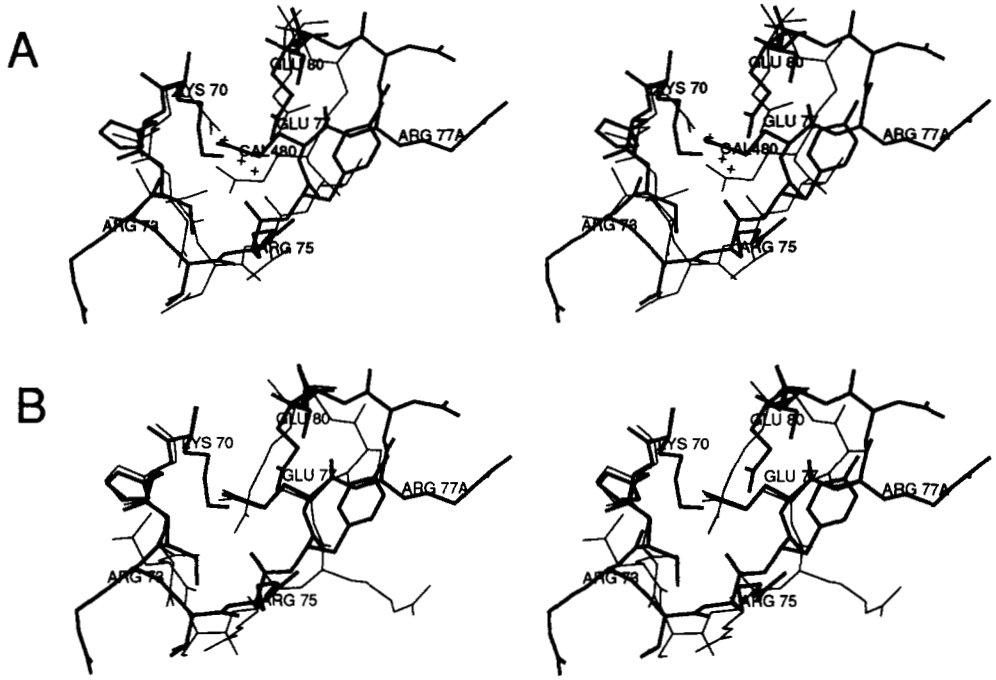

Fig. 15. A: Loop Lys 70–Glu 80 of human α-thrombin (thick connections) superimposed by the calcium-binding loop, the bound calcium ion, and two adjacent solvent molecules (crosses) of bovine trypsin (thin connections [Bode & Schwager, 1975a,b]). Nζ of Lys 70 of thrombin occupies an equivalent position as the calcium ion in trypsin. The view is similar to the standard view of Figure 1. **B**: Loop Lys 70–Glu 80 of human α-thrombin (thick connections) superimposed by the equivalent loop of human leukocyte elastase (thin connections [Bode et al., 1986]). The guanidinium group of Arg 80 of the elastase occupies a similar position as Lys 70Nζ of thrombin.

the static α-thrombin structure (Fig. 14B), is (in the absence of shielding proteins such as hirudin) highly susceptible to chemical modification (Chang, 1989; Church et al., 1989). Clearly, ease and amount of modification are not necessarily reliable indicators for surface accessibility but might correlate with other molecular properties.

In turn, residue Glu 80 is involved in another salt bridge/hydrogen bond with the guanidyl group of Arg 67. Thus, the charged groups of four residues (Arg 67, Lys 70, Glu 77, and Glu 80) are cross-connected with one another to form a salt-bridge cluster essentially buried in the molecule (Fig. 14A). Its contribution to the rigidity and integrity of this loop appears considerable (see Table 6). As shown previously for trypsin (Bode, 1979), the integrity of this loop structure confers (thermal) stability to the whole molecule. Thrombin (like porcine kallikrein [Bode et al., 1983] and human leukocyte elastase [Bode et al., 1986; Fig. 15B; Table 11]) seems to carry its endogenous cationic group (70Nζ) to order the 70–80 loop and to maintain a distinct surface contour (Fig. 14A,B); thus, these three proteinases share similar 70–80 loops with respect to conformation and a central basic residue.

As already discussed (Table 7) and shown further in Figure 14B, the slightly notched surface arching over this loop has only positively charged amino acid side chains; in particular Arg 73, Arg 75, and Arg 77A are surrounded exclusively by other positively charged residues such as Arg 35, Lys 149E, Lys 81, Lys 110, Lys 109, and Lys 36, which are themselves quite distant to more peripheral negatively charged residues on the surface. Their positive charge is partially compensated by the negative charges of Glu 77 and Glu 80 involved in the ionic cluster beneath the surface (see above and Fig. 14); nevertheless, they give rise to a strong positive field around this surface (Fig. 8; Table 7).

The crystal structure analysis of the thrombin–hirudin complex (Rydel et al., 1990) revealed that this positively charged thrombin surface patch interacts with the extended hirudin tail segment (Fig. 16). Beside these distinct electrostatic interactions of hirudin, the complementary fit of hydrophobic groups seems to be of particular importance (see also Stone et al., 1987; Mao et al., 1988; Maraganore et al., 1990). There is much evidence that this thrombin exosite is also involved in thrombin binding to fibrinogen (Hageman & Scheraga, 1974; Van Nispen et al., 1977; Hogg & Blombäck, 1978; White et al., 1981; Hofsteenge & Stone, 1987; Fenton et al., 1988; Lundblad et al., 1988; Church et al., 1989; Hogg & Jackson, 1989), to fibrin and its E-domain (Liu et al., 1979; Berliner et al., 1985; Kaminski & McDonagh, 1987; Kaczmarek & McDonagh, 1988; Vali & Scheraga, 1988), to thrombomodulin (Hofsteenge et al., 1986; Hofsteenge &

Table 11. *Distances between the central cationic group and the ligands in the crystal structures of human α-thrombin (HTHR), bovine trypsin (BTRY; entry DEBA in the Protein Data Bank [Bode & Schwager, 1975a,b; Bernstein et al., 1977], human leukocyte elastase (HLE) [Bode et al., 1986], and porcine pancreatic kallikrein (KK) [Bode et al., 1983]*

Central cationic group	Liganding group		Distance (Å)
HTHR			
Lys 70Nζ	..... Arg 75	O	2.6
	..... Glu 77	OE2	2.6
	..... Glu 80	OE1	2.7
BTRY			
Ca^{2+}	..... Glu 70	OE1	2.3
	..... Asn 72	O	2.3
	..... Val 75	O	2.3
	..... Glu 80	OE(1)	2.3
	..... Sol 711	OH	2.3
	..... Sol 714 (with connection to Glu 77)	OH	2.3
HLE			
Arg 80NE	..... Glu 77	OE	3.1
NH2	..... Asn 72	O	3.1
	..... Arg 75	O	2.9
KK			
Arg 70 guanidyl group	 72	O	–
(weakly defined)	 75	O	–
	..... Glu 77	OE	–

Stone; 1987; Preissner et al., 1987; Suzuki et al., 1990), and to anionic matrices (see, e.g., Fenton et al., 1988, 1989). Very recently, it has been shown that this exosite is also involved in interaction with a negatively charged peptide segment of the platelet-derived thrombin receptor (Liu et al., 1991).

It has been shown that the above proteins bind competitively (Hofsteenge et al., 1986), i.e., that they probably occupy similar patches of this positively charged exosite of thrombin. Segment 30–44 of the fibrinogen Aα-chain, which comprises several negatively charged residues, seems to be particularly important for binding and acceleration of the activation cleavage (Blombäck et al., 1977; Van Nispen et al., 1977; Hofsteenge et al., 1986). This fibrinogen segment exhibits only a very weak sequence homology with the hirudin tail segment, however. The thrombin–hirudin complex (Fig. 16) can, nevertheless, serve as a model for the interaction of thrombin with fibrinogen.

The positive field centered at this exosite (Fig. 8) spreads far into the extramolecular space; consequently, approaching macromolecules (substrates as well as inhibitors) could be preoriented by the combined influence of this basic exosite and the negatively charged patch around the specificity pocket (see below). Anionic groups of associated proteins need not necessarily interact through direct salt bridges with the cationic thrombin groups; they nevertheless experience the effect of this positive field at some distance from the thrombin surface. Due to its interaction with various anionic compounds and protein segments, this exosite is usually referred to as the anion-binding exosite (Fenton, 1986). This could, however, lead to confusion with the other principal basic site described below. More appropriate designations would therefore be the terms anion-binding exosite I or fibrin(ogen) secondary binding exosite (Fenton, 1981), reflecting the importance of this exosite in fibrinogen clotting and fibrin binding.

The exposed peptide bond Arg 77A–Asn 78, in the 70–

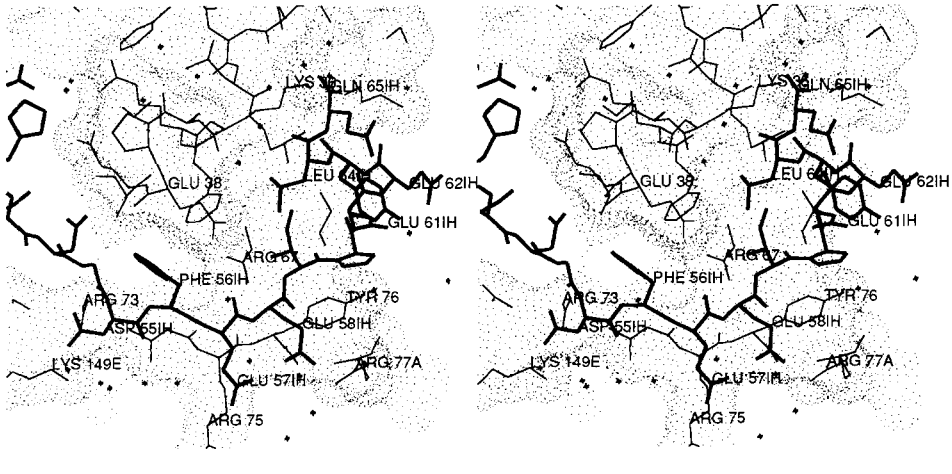

Fig. 16. The fibrinogen secondary binding exosite of thrombin (thin connections) and the carboxy-terminal hirudin segment (thick connections) of thrombin–hirudin (Rydel et al., 1991), displayed together with the Connolly surface of thrombin. The view is similar to the standard view of Figure 1.

80 loop (Fig. 14A), is particularly susceptible to tryptic and to autocatalytic attack (Fenton et al., 1977; Boissel et al., 1984; Chang, 1986; Bezeaud & Guillin, 1988) giving rise to β_T-thrombin (Braun et al., 1988). This polypeptide segment is relatively mobile (in particular from Thr 74 to Ile 79; see Fig. 28). Its main-chain conformation is quite different from the canonical conformation of typical serine proteinase protein inhibitor binding loops (Huber & Bode, 1978; Laskowski & Kato, 1980; Read & James, 1986; Bode & Huber, 1991) (see Table 10). Human β-thrombin is formed upon its cleavage and following excision of peptide Ile 68–Arg 77A.

Experimental evidence (Noé et al., 1988) suggests that the whole loop structure unfolds upon cleavage, driven presumably by favorable interactions of charged residues (previously involved in the buried salt cluster) with bulk water. Such a disruption of the exosite surface would explain the tremendous reduction in clotting activity and loss of affinity for fibrin, thrombomodulin, and hirudin, observed for β_T-thrombin (Braun et al., 1988; Hofsteenge et al., 1988) and β-thrombin (Fenton et al., 1977; Hofsteenge & Stone, 1987; Lewis et al., 1987; Stone et al., 1987; Bezeaud & Guillin, 1988). Indirect evidence for such an unfolding is that peptide bond Arg 67–Ile 68 (Fig. 14A) is rather susceptible to autocatalytic cleavage (Boissel et al., 1984). Being deeply buried, a cleavage of the surface Arg 77A–Asn 78 bond (see Fig. 14A,B) would be necessary prior to a proteolytic attack at this site to give β-thrombin. As witnessed by the almost unchanged kinetic parameters of β-thrombin toward chromogenic substrates this loop and surface disintegration does not seem to affect the geometry at the active site.

In the naturally occurring human genetic variant Quick I, Arg 67 (see Fig. 12) is replaced by a cysteine residue (Henriksen & Mann, 1988). The surface location of the Arg 67 guanidyl group (Fig. 14B) and its integration in an extended internal salt-bridge cluster in normal human α-thrombin suggests that the exosite of this mutant might be disrupted in a similar manner as in β-thrombin. This agrees with the experimental observations of tremendous loss in clotting activity but with retention of hydrolytic activity toward small synthetic substrates; the considerable reduction of platelet stimulation capability found for this mutant is a hint that this exosite is involved in platelet aggregation.

Insertion loop Leu 144–Gly 150 (149 insertion loop)

The polypeptide segment between residues 144 and 150 is five amino acid residues longer in thrombin than in chymotrypsin (see Table 3). The exposed loop (Fig. 17) lacks any inter-main-chain hydrogen bonds, but seems to be stabilized through hydrogen bonds made with its own side chains (Lys 145 and Asn 149B); some hydrophobic contacts with a symmetry-related molecule seem to further stabilize this loop in the PPACK–thrombin crystals. This loop segment attains a very different conformation and overall shape in thrombin–hirudin (Fig. 17) and is of variable conformation in the various bovine thrombin structures. In the PPACK–thrombin conformation, the Trp 148 side chain would collide with parts of the bound hirudin molecule (Rydel et al., 1991). The loop conformation observed in PPACK–thrombin may be enforced by crystal or inhibitor contacts. There is considerable evidence (e.g., from the bovine thrombin crystals) that this loop exhibits a significant degree of overall flexibility.

Three proteolyzed human thrombin variants have been found and characterized cleaved in this exposed, flexible loop segment (displayed in Fig. 12), namely (1) ξ-thrombin, resulting from cathepsin G cleavage at Trp 148–Thr 149 (Brezniak et al., 1990); (2) ϵ-thrombin, generated by elastase action on Ala 149A–Asn 149B (Kawabata et al., 1985; Brower et al., 1987); and (3) γ-thrombin, resulting from tryptic or autocatalytic cleavage at Lys 149E–Gly 150 (Bing et al., 1977; Boissel et al., 1984).

As this loop is inherently flexible, the first two cleav-

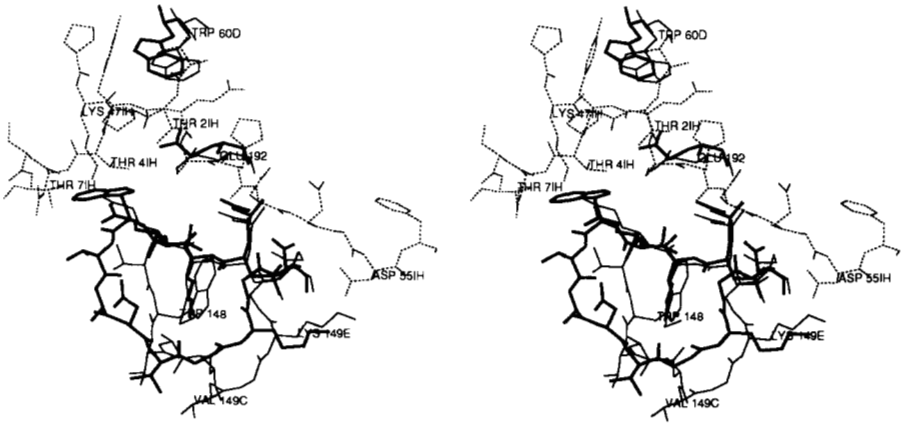

Fig. 17. Segment Asn 143–Gln 151 and Glu 192 of PPACK–thrombin (thick connections) superimposed with the same residues of the thrombin component (thin connections) and adjacent segments of hirudin (dashed connections) of the thrombin–hirudin complex (Rydel et al., 1991). Thrombin residue Trp 148 is very differently located in both thrombin molecules, and loop Glu 146–Gly 150 is differently arranged.

ages should not greatly affect the active-site geometry and thus the catalytic activity of thrombin, as demonstrated experimentally (Stone et al., 1987; Hofsteenge et al., 1988). Nevertheless, an intact loop might confer some extra stability to the whole molecule.

The last of these scissile peptide bonds, at Lys 149E–Gly 150 (Fig. 17), is somewhat obstructed in PPACK-α-thrombin and not readily accessible to attacking protease molecules. A primary autolytic cleavage seems possible at this site, however (Chang, 1986), probably due to the inherent mobility of this loop. This scissile peptide bond is adjacent to segment Gln 151–Pro 152, which forms (together with Glu 39 of the opposing loop) the entrance to the fibrin(ogen) secondary binding exosite (see above and Fig. 14B). ϵ-Thrombin, cleaved four bonds before Lys 149E–Gly 150 (see Fig. 12), shows only a minor decrease in clotting activity compared with α-thrombin (Hofsteenge et al., 1988). The nearly complete lack of clotting activity of γ_T- and γ-thrombin (Bing et al., 1977; Fenton et al., 1977) might therefore be caused mainly by the disruption of this exosite due to excision of peptide Ile 68–Arg 77A (Hofsteenge et al., 1988), i.e., similar to β-thrombin.

Peptide segment Thr 147–Ser 153 might be involved in thrombomodulin binding (Suzuki et al., 1990). Other evidence (Hofsteenge et al., 1986; Preissner et al., 1987; Noé et al., 1988; Jakubowski & Owen, 1989; Tsiang et al., 1990) indicates, however, that the accessibility and integrity of the 70–80 surface loop and the fibrin(ogen) secondary binding exosite are at least also required for efficient thrombin interaction with thrombomodulin. It has been shown (LeBonniec & Esmon, 1991) that a Glu 192 → Gln substitution in thrombin mimics the catalytic switch induced by thrombomodulin. Thus, the thrombin interaction site for this endothelial cell-surface protein seems to comprise a larger surface area including the junction between loop segments 70–80 and 144–151 (Wu et al., 1991).

The putative heparin-binding site

A thrombin surface patch of positively charged side chains (Arg 126, Lys 236, Lys 240, and Arg 93) surrounded by other basic residues (Arg 101, Arg 233, Arg 165, Lys 169, Lys 235, Arg 175, Arg 173, Arg 97, patch p2 in Table 7), which in turn contact negatively charged side chains (see also Tables 5, 6), is shown in Figure 18A. These positively charged residues are quite densely packed on the surface and are not compensated by negatively charged groups below the surface as for the 70–80 loop. Each of them therefore gives rise to a very strong positive electrostatic field (Fig. 8). In spite of this strong field, we see no evidence for fixed solvent anions (Fig. 18B); the isolated blobs of electron density within the surface depression and in the intermolecular (crystal) gap are sufficiently interpreted as water molecules. This site seems rather to be more suited to binding polyanions such as heparin; this glycosaminoglycan of $1\frac{1}{2}$ acidic (sulfato and uronic) groups per monosaccharide unit binds to thrombin (Griffith et al., 1979) and accelerates its inactivation by antithrombin III (Björk & Lindahl, 1982), presumably through interactions with alkaline groups of thrombin and antithrombin.

Chemical modification experiments on thrombin revealed that Lys 169 and Lys 240 (at the edges of this positively charged patch [Fig. 18A]) are blocked by bound heparin and that their modification renders thrombin inaccessible to heparin (Church et al., 1989). Cooperative electrostatic interactions of linear heparin molecules (with high linear negative charge density) with this positively charged region of contiguous basic residues could explain the high affinity for binding heparin to thrombin (Heuck et al., 1985). Furthermore, this putative heparin-binding site exhibits a typical heparin-binding motif (a helical strand comprising several basic residues [see Table 3; Cardin & Weintraub, 1989]). Some slippage of this thrombin helix caused by heparin binding cannot be excluded bearing in mind the relatively variable arrangement of the carboxy-terminal helix in related serine proteinase-reactive domains (Bode & Huber, 1986).

Acceleration of complex formation between antithrombin III and thrombin by heparin may be due in part to a template mechanism (Pomeranz & Owen, 1978; Griffith, 1982; Nesheim, 1983) in addition to its effects on the antithrombin III structure. If a docking geometry to be similar to that known for the "smaller" protein inhibitors (Bode & Huber, 1991) is assumed for antithrombin III bound to thrombin, the thrombin heparin-binding site would be located near the presumed heparin-binding site of antithrombin III in antithrombin III–thrombin complexes (for references, see Huber & Carrell, 1989). It is therefore conceivable that one heparin strand of about 14 monosaccharide units (required for efficient acceleration of the thrombin–antithrombin III reaction [see Huber & Carrell, 1989]) could connect the two components. The Arg 97 and Arg 175 residues, which form a positively charged satellite patch (p2c in Table 7; Figs. 7, 8, 18) to the "north rim" (Bode et al., 1989b) of the active-site canyon could further stabilize such a mediatory binding through heparin.

Similar reaction rates of antithrombin III with α- and with γ-thrombin in the presence of heparin support the idea that the main heparin recognition sites of antithrombin III–bound thrombin reside outside the fibrin(ogen) secondary binding exosite (Stone & Hofsteenge, 1987). Other evidence (Olson et al., 1986; Hogg & Jackson, 1989; Naski et al., 1990) indicates, however, that heparin might bind to both exosites in the absence of antithrombin III. This highly charged putative heparin-binding site might also (in cooperation with the fibrin(ogen) secondary binding exosite) mediate thrombin binding to cation exchange resins and fibrin (Fenton et al., 1988) and might therefore be alternatively called anion-binding exosite II.

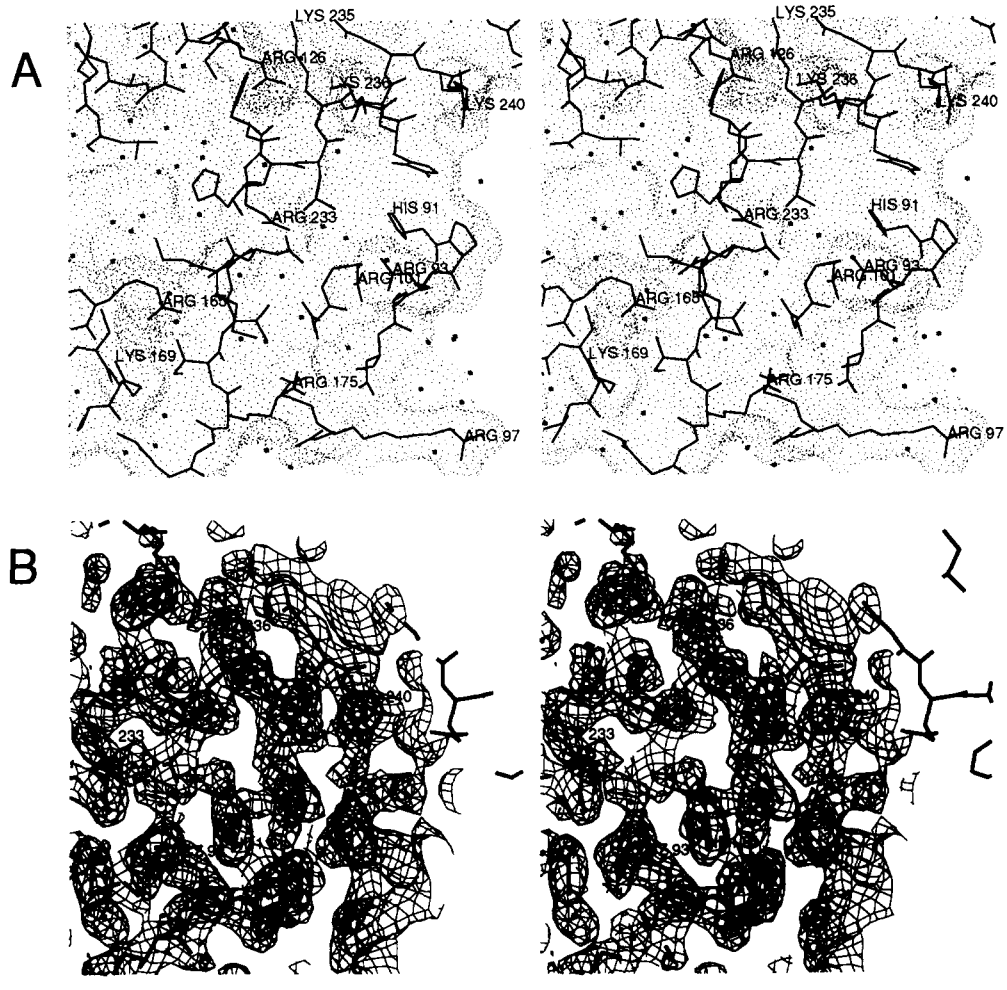

Fig. 18. Putative heparin-binding site of human α-thrombin. The view is obtained from the standard view upon an approximately 90° rotation around a horizontal axis. **A**: Surface structure between Lys 169 (lower left corner) and Lys 240 (upper right corner), displayed together with the Connolly surface. The site of highest positive electrostatic potential is in the surface depression close to His 91, surrounded by Arg 126, Lys 236, Lys 240, Arg 93, Arg 101, and Arg 233. Crosses represent localized solvent molecules. **B**: Section of the final electron density superimposed with equivalent surface parts of thrombin and a few localized solvent molecules (crosses). The depression outside the thrombin surface does not contain appropriate high density, which could account for a bound anion. Contour surface at 0.8σ.

Furthermore, the heparin-binding site is close to the amino-terminal Thr 1H of the A-chain (see Fig. 7); its partial shielding by adjacent pro-parts in prothrombin (Fenton et al., 1988) or in other partially activated prothrombin forms still possessing pro-parts (such as meizothrombin [see Doyle & Mann, 1990]) also seems possible therefore without conformational changes in the B-chain globule. Chemical modification experiments further suggest that part of this heparin-binding site (in particular Lys 169 and/or Lys 240) might act simultaneously as a high-affinity binding site for platelets (White et al., 1981).

In the natural human variant thrombin Tokushima, Arg 101 is replaced by a tryptophan residue (Miyata et al., 1987). The side chain of Arg 101 points away from the active-site cleft (Fig. 12) but is involved with the adjacent Asp 100 in a surface-located salt bridge (Table 6). According to modeling experiments, substitution with a tryptophan side chain might disturb the "rim" region and would thus primarily affect binding of more extended substrates and interactions with cells. An effect on the subsequent active-site residue, Asp 102, whose side chain points in the opposite direction (toward the active-site cleft [see Fig. 1]), cannot be excluded; it would be small, however.

The thrombin RGD site

Both human and bovine thrombin contain an RGD segment (Arg 187–Gly 188–Asp 189) beneath the specificity pocket (Fig. 19). This segment has recently been suggested (Fenton, 1988) to play a distinct role as a docking

site in platelet binding (e.g., through the platelet-specific GPIIb/IIIa integrin receptor [see Charo et al., 1987]). Glenn et al. (1988) showed that a thrombin-derived 23-amino acid peptide comprising this Arg-Gly-Asp region competes with α-thrombin for binding to the high-affinity receptor on fibroblasts, and that it favors the thrombin-mediated thymidine incorporation; in contrast, the RGDA peptide, which likewise competes with α-thrombin, inhibits the mitogenic effect of α-thrombin. Bar-Shavit et al. (1991) demonstrated recently that the endothelial cell adhesion properties of modified human thrombin could be inhibited specifically by a GRGDSP peptide; antithrombin III inhibited this thrombin adhesion, whereas hirudin did not interfere with it.

In the PPACK–α-thrombin structure, the side chain of Arg 187 runs along the molecular surface; it is exposed to solvent with the distal guanidyl group engaged in quite strong salt/hydrogen-bond bridges (see Fig. 19; Table 6). Neither Gly 188 nor Asp 189 (except for its carboxylate group, which points toward the interior of the specificity pocket [Fig. 19]) are, however, accessible to receptor structures of approaching cells or proteins without complete unfolding of this thrombin site. Such an accessibility would seem to be required, however, for proper recognition of the underlying sequence motif. In partially degraded or thermally unfolded thrombin derivatives or in zymogenlike (pro)thrombin species (with the Asp 189-containing peptide fragment turned inside-out as in chymotrypsinogen, see above) a greater exposure of this RGD site is conceivable; this would be in agreement with the observation that intact α-thrombin attains endothelial cell adhesive properties only after exposure to elevated temperatures (Bar-Shavit et al., 1991).

The active-site cleft

The most remarkable feature of the thrombin surface is a deep narrow "canyon" (see Fig. 1) containing the catalytic residues and the adjacent substrate-binding subsites (Bode et al., 1989b). The catalytic residues (in particular Ser 195, His 57, and Asp 102) divide this cleft into two halves, corresponding to a fibrinopeptide side ("west") and a fibrin side ("east"). This cleft is bordered and shaped by the four extending polypeptide segments, Arg 173–Ile 174, Asp 95–Leu 99, Tyr 60A–Phe 60H, and Phe 34–Cys 42, on one side (forming the "north rim," Fig. 1) and the two loop segments, Gly 216–Cys 220 and Asn 143–Gln 151, on the opposite side (the "south rim"). The side chains of Ile 174, Tyr 60A, Trp 60D, Lys 60F, Phe 60H, and of Thr 147, Trp 148, respectively, protrude into the active-site cleft, i.e., each rim is lined primarily by hydrophobic groups. The base, in contrast, is covered mainly with polar/charged side chains (Glu 217, Glu 192, Asn 143, Gln 151, Arg 73) and polar main-chain groups.

Around the entrance to the specificity pocket (Fig. 20), 10 negatively charged residues (among them Asp 189, Glu 192, Asp 221, and Glu 146) are clustered (patch n1a in Table 7), with Glu 192 (relatively flexible at the base) and Asp 189 (at the bottom of the specificity pocket [see Fig. 20]) being the only acidic residues not involved in direct salt bridges. The latter two, together with Asp 221, Glu 18, Asp 186A, and Glu 186B, exert repulsive forces on one another and contribute substantially to the rather negative field generated at this site (see Fig. 8; Tables 5, 7). According to our calculations, the catalytic residues are only marginally affected by this negative electric potential (due to their position at the edge of the negative

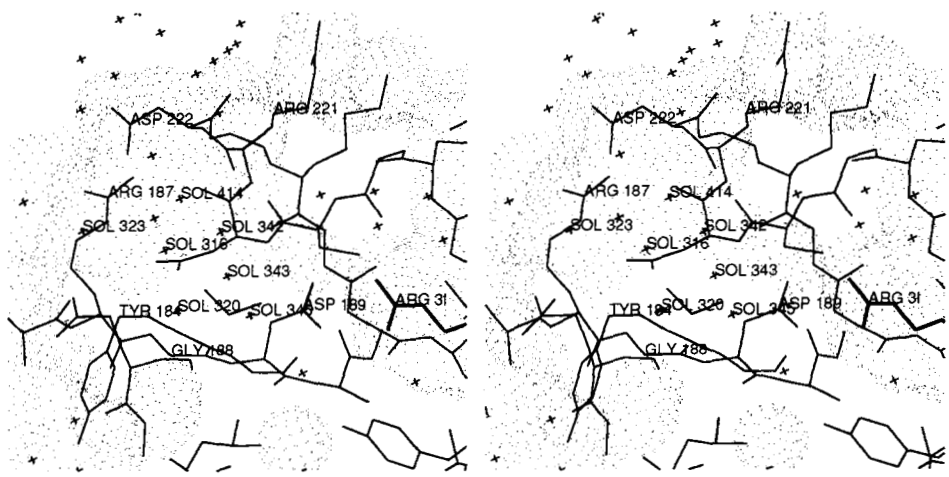

Fig. 19. Segment Arg 187–Gly 188–Asp 189 and adjacent residues of human α-thrombin (thin connections), displayed together with the Connolly surface. Most of the internal solvents (crosses), which form an extended solvent cluster, are labeled. Arg 187 is located at the thrombin surface; Gly 188 and Asp 189 are, however, buried in the interior of the thrombin molecule. The carboxylate group of Asp 189 is directed toward the specificity pocket and in ion pair contact with the guanidyl group of Arg 31 of PPACK (thick connections). The main-chain angles of this segment are −138°, 174° (Arg 187); 132°, 155° (Gly 188); −171°, 16° (Asp 189). The view is similar to the standard view of Figure 1.

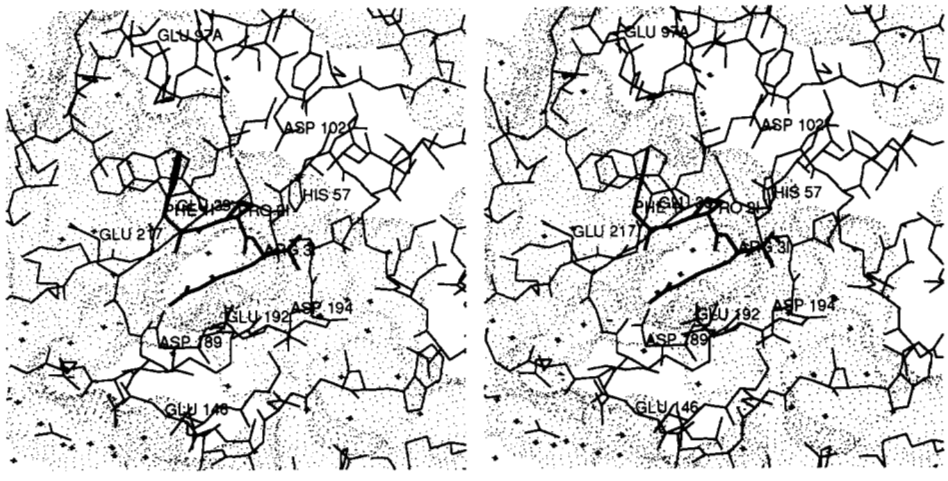

Fig. 20. The active-site region of PPACK–thrombin, displayed together with the Connolly surface of thrombin. The thrombin structure is given with thin connections, the PPACK molecule with thick connections, and localized solvent molecules are represented by crosses. The 60 insertion loop, which partially occludes the active-site residues and the entrance to the specificity pocket, is cut off; thus, the active-site residues and Pro 2I and Arg 3I are freely visible. The negatively charged residues arranged around the specificity pocket to contribute to the negative patch n1a (Table 7) are labeled. Standard view as in Figure 1.

charge-patch); besides their mutual interactions, they experience mainly the ammonium group of Lys 60F. The latter interaction should shift the pK values of Asp 102 and His 57 to slightly lower values than normal. As described above, the fibrin(ogen) secondary binding exosite (at the right-hand side of the active-site cleft [Figs. 1, 14]) in contrast gives rise to a strong positive electrostatic field (see Fig. 8). This dipolar charge distribution along the active-site cleft would clearly influence the orientation of approaching substrates and inhibitors of large electric moments (such as hirudin, see above).

The thrombin active site and interaction with PPACK

The PPACK molecule, as it is bound to the active-site cleft of thrombin, is shown in Figure 21 and Kinemage 6. Its peptidyl moiety juxtaposes the extended thrombin segment Ser 214–Glu 217 in a twisted antiparallel manner. The carboxy-terminal Arg 3I carbonyl group is part of a tetrahedral hemiketal structure; i.e., it is linked covalently to Ser 195 Oγ and, via the methylene group, to the imidazole group (Nε2) of His 57 of throm-

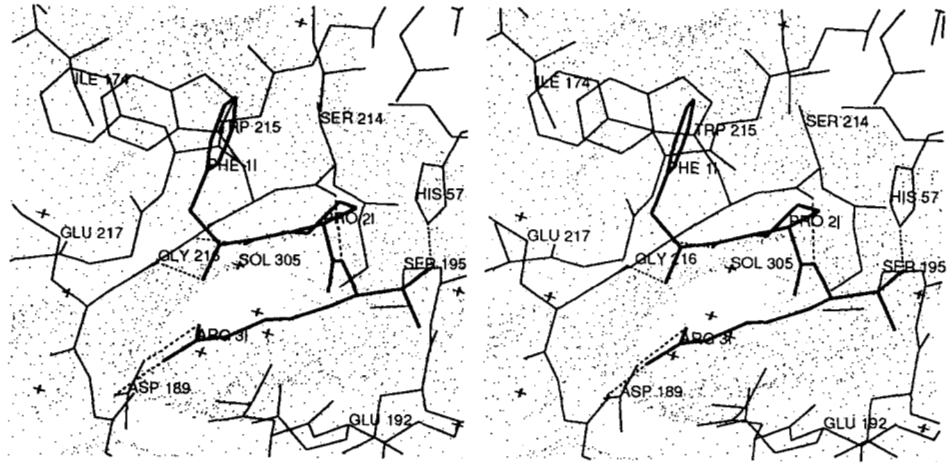

Fig. 21. Interaction of the PPACK molecule (thick connections) with the thrombin active site (thin connections), displayed together with the Connolly surface of thrombin and some localized solvent molecules (crosses). Most of the hydrogen bonds formed between PPACK and thrombin are shown in addition (dashed lines). The Arg 3I carbonyl group forms a tetrahedral hemiketal structure covalently linked with Ser 195Oγ and (via the methylene group) with His 57Nε2. Standard view as in Figure 1.

bin. The side chain of Ser 195 is probably considerably rotated to form this covalent bond compared to its location in unliganded thrombin.

The thrombin specificity pocket is geometrically quite similar to that of bovine trypsin (Fig. 22) (Bode & Schwager, 1975a). The r.m.s. deviation of 22 α-carbon atoms of the residues and segments lining the specificity pocket in PPACK-thrombin is 0.39 Å compared with bovine trypsin; an almost identical deviation is obtained upon comparison of PPACK-thrombin with hirudin-thrombin. The only thrombin residue whose side chain points into the pocket differently is Ala 190, replacing a serine residue in trypsin. This exchange renders the thrombin pocket slightly less polar but, in particular, causes the loss of a hydrogen-bond acceptor/donor (perpendicular to the guanidyl group of Arg 3I [Fig. 22]). Solvent molecule 305, which becomes buried upon complex formation, is conserved in the thrombin pocket (compare with trypsin [Bode & Schwager, 1975a]). This solvent molecule finds a much more hydrophobic surrounding in thrombin than in trypsin, but is nevertheless adequately positioned to receive a favorable hydrogen bond from the NH1 atom of the Arg 3I guanidinium group. In consequence, the thrombin S1 subsite would better accommodate an arginine P1 side chain, whereas an inserted lysine side chain would fit less tightly in the thrombin pocket, due to the greater space and lack of a hydrogen-bond partner (residue 190). This is in agreement with experimental results concerning the relative affinities for arginine and lysine P1 side chains (Liem & Scheraga, 1974; Kettner & Shaw, 1979, 1981; Lottenberg et al., 1983).

The amidino nitrogen of the P3 residue Arg 3I is 2.9 Å away from Ser 214 O, and in a relatively favorable orientation to form a hydrogen bond (Fig. 21). Formation of this hydrogen bond is assumed to be an important step to achieve the tetrahedral transition state in productive enzyme-substrate complexes; thus, the stable state seen here mimics somewhat a real transition state.

Quick II (Henriksen & Mann, 1989), another naturally occurring dysfunctional thrombin mutant, has the replacement Gly 226 → Val 226 (see Fig. 12). Modeling experiments show that the bulky valine side chain would probably disturb the Asp 189 conformation at the bottom of the pocket, rather than directly blocking binding of an inserted arginine side chain. The pocket of Quick II will probably not be as narrowed as in porcine pancreatic elastase (Meyer et al., 1988) or human leukocyte elastase (Bode et al., 1986). Thus, neither typical trypsin nor elastase substrates should be cleaved efficiently by this mutant, in agreement with the experimentally observed lack of catalytic and clotting activity (Henriksen & Mann, 1989).

Behind Asp 189, the thrombin insertion loop Tyr 184A–Gly 188 and segment Asp 221A–Tyr 225 enclose a large internal cavity filled with (buried) solvent molecules (see Fig. 19; Table 9). This thrombin water cluster is even larger than in trypsin, due to the smaller Thr 172 of thrombin compared with a tyrosine in trypsin. This cluster and its linkage to the bulk water have previously been suggested as providing a route to expel water molecules through the back of the specificity pocket upon insertion of a P1 side chain of a binding peptide (Meyer et al., 1988).

Another more extended, quite hydrophobic pocket lies near the entrance to the specificity pocket (Fig. 21), lined by Ile 174, Trp 215, segment 97–99, His 57, Tyr 60A, and Trp 60D. This hydrophobic depression is most certainly the suggested apolar binding site (Berliner & Shen, 1977). Its back part is made up of the same structural elements as in trypsin (Trp 215, a shorter segment 97–99, His 57 [see Fig. 22]); it is in thrombin, however, further screened off from bulk water by the mainly hydrophobic Tyr 60A–Trp 60D loop on one side and Ile 174 on the other side (to the right and left in Fig. 21, respectively). The PPACK molecule nestles tightly into this depression. The solvent-accessible surface of the PPACK–thrombin complex is 240 Å^2 smaller than that of thrombin itself; only one-third of the PPACK surface (one edge of the benzene ring of D-Phe 1I, and Pro 2I [see Fig. 21]) does not become buried upon the complex formation. In the PPACK–thrombin complex the pyrrolidine ring of Pro 2I is almost fully shielded from bulk water; it is in a completely hy-

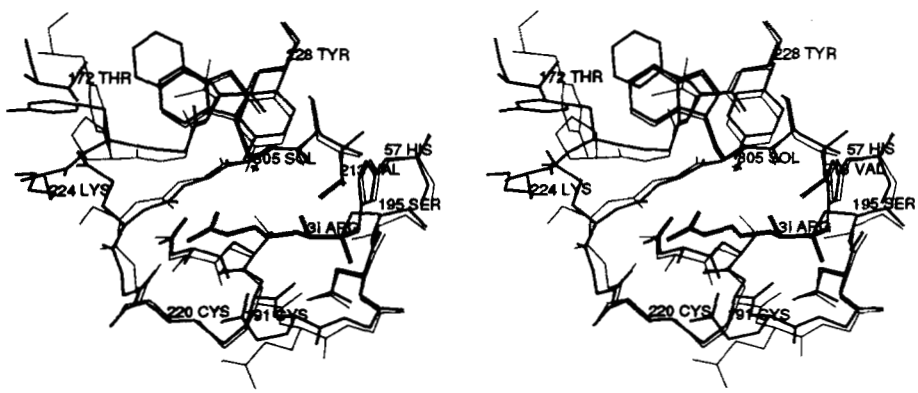

Fig. 22. Human α-thrombin structure around the specificity pocket (medium connections) superimposed by the equivalent residues of unliganded bovine trypsin (thin connections [Bode et al., 1976]). Arg 3I of PPACK is shown with thick connections. The view is into the specificity pocket, i.e., similar to the standard view of Figure 1. Ser 190 of trypsin is replaced in thrombin by an alanine residue. Sol 305 in thrombin has an equivalent water molecule (Sol 416) in trypsin.

drophobic environment and seems to be perfectly accommodated through its polyproline II conformation (see Fig. 21). This excellent packing in the thrombin S2 cavity explains the frequent occurrence of proline (and other medium-sized hydrophobic) residues at the P2 position of naturally occurring thrombin macromolecular substrates. In contrast to trypsin (Liem & Scheraga, 1974; Bajusz et al., 1978), it further accounts in thrombin for the particular reactivity of chloromethylketone inhibitors with P2-proline residues (Kettner & Shaw, 1981). Much larger residues would not pack favorably into the S2-thrombin cavity, and smaller ones (such as glycine residues) would not be as well fixed.

The (substituted) piperidine rings of the benzamidine (NAPAP [Stürzebecher et al., 1983]) and of the arginine-based (MQPA [Okamoto et al., 1981]) synthetic inhibitors (Fig. 23) occupy this pocket where they can be tightly fixed and sandwiched between the flat side of the His 57 imidazole ring and the amino-terminal aryl moiety (Bode et al., 1990; Turk et al., 1991; Brandstetter et al., 1992). As in the case of PPACK-thrombin, the hydrogen bonds made with Gly 216 N and O determine mainly the docking position of these inhibitors in the thrombin active site.

The side chain of the preceding residue of PPACK, D-Phe 1I, fits into the notched, mainly hydrophobic cleft made by Ile 174, Trp 215, segment 97–99, and Tyr 60H (Fig. 21). Such an interaction is not only preferred due to favorable hydrophobic contacts; presumably, both the perpendicular aryl–aryl (edge-on) arrangement (Burley & Petsko, 1985) and the aryl–carbonyl contact with the carbonyl group of Glu 97A (see Thomas et al., 1982) contribute significantly to binding strength (Bode et al., 1990). In natural protein substrates, the variety of amino acids in P3 is great; in general, L-amino acids in P3 position would interact with a different thrombin site. The P4 side chain of a bound all-L-amino acid polypeptide substrate would, however, point toward this hydrophobic cleft; in fact, the beneficial effect of large bulky aromatic residues at P4 connected to a P3 glycine has been used to construct suitable thrombin substrates (Svendsen et al., 1972). Very recently we have shown that the naphthyl and tosyl moieties of certain benzamidine-derived synthetic inhibitors (see Fig. 23; Bode et al., 1990; Turk et al., 1991) as well as other aromatic groups in arginine-derived synthetic inhibitors (Matsuzaki et al., 1989) interact in a similar favorable manner with this cleft (Brandstetter et al., 1992). The phenolic side chain of Tyr 3HI of hirudin is also placed in such a manner (Rydel et al., 1991), and there is now crystallographic evidence (Martin et al., 1992; Stubbs et al., 1992) for an equivalent arrangement of the phenyl moiety of the fibrinogen Aα-chain Phe 8 residue upon binding to thrombin similar to previous suggestions (Meinwald et al., 1980; Ni et al., 1989a,b). We have therefore proposed the designation aryl-binding site (Bode et al., 1990), distinguishing it from the other part of the apolar binding site (the S2 cavity).

The extraordinary specificity and reactivity of PPACK for thrombin (Kettner & Shaw, 1979) can thus easily be attributed to the optimal fit of all three residues to their respective thrombin subsites (see Fig. 21). In addition, each of these PPACK residues seems to exist in a stereochemically and energetically optimal conformation in the thrombin-bound PPACK molecule (see Table 10): both the main- and the side-chain moieties of Arg 3I are arranged in an extended conformation of favorable energy; Pro 2I exhibits a characteristic polyproline II conforma-

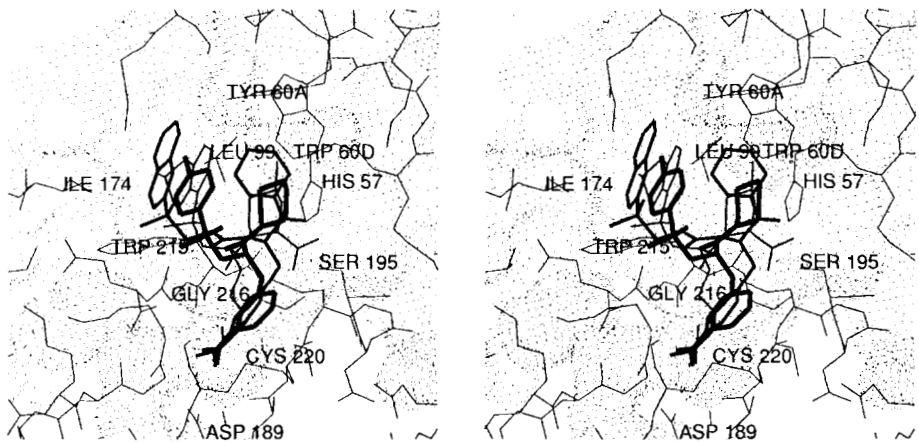

Fig. 23. Tosyl-*m*-amidinophenylalanyl-piperidine (thick connections), NAPAP (medium connections), and MQPA (thin connections) bound to the active site of human α-thrombin displayed together with the Connolly surface of thrombin (Turk et al., 1991). The naphthyl/toluene/chinolyl groups of the inhibitors interact with the aryl-binding site of thrombin; the side chains of the *m*- and the *p*-amidinophenylalanyl residues and of the arginyl residue enter the specificity pocket from slightly differing sites; the S2 subsite of thrombin is occupied to different extents by the (partially substituted) piperidine moieties. The view is similar to the standard view of Figure 1.

tion; and D-Phe 1I (with a Cα-C angle of −131°) represents just one of the two optimal conformers preferred for steric reasons by a D-amino acid residue linked to a proline.

The ammonium and the carbonyl group of D-Phe 1I are in favorable hydrogen-bond contacts with Gly 216 (2.7 Å [see Fig. 21]). In contrast to trypsin, good peptidyl aldehyde inhibitors of thrombin require termination at P3 (Bajusz et al., 1978), probably to allow the undisturbed formation of this hydrogen bond; it is clear, however, that *N*-alkylation of PPACK (such as in *N*-methyl-PPACK [see Bajusz et al., 1990]) should not impair hydrogen-bond formation in solution. In the PPACK–thrombin crystal structure, the ammonium group forms two further hydrogen bonds with a fixed solvent molecule and with Gln 244 of a symmetry-related thrombin molecule. This contact might confer stability to this PPACK–thrombin crystal form.

An L-configured Phe 1I residue at P3 would not be able to interact with thrombin in as favorable a manner as the D-diastereomer (in particular when linked to a proline [see Fig. 21]). Simultaneous hydrogen-bond formation through its amino group and favorable aryl interaction are mutually exclusive; the best fit for an L-Phe 1I residue would seem to be at a Cα-C dihedral angle of about −60°, with the phenyl group fitting into the aryl-binding site in a slightly different relative orientation.

Due to lack of direct structural evidence, the primed subsites of thrombin cannot be described with a confidence equivalent to the nonprimed sites. The thrombin complexes of various protein inhibitors (see, e.g., Fig. 24), modeled according to experimental serine proteinase complexes, might give a rough idea of the probable interactions of substrate P1′, P2′, and P3′ residues with thrombin. The S1′ subsite of thrombin is (due to the impending Lys 60F) limited in size and therefore particularly suited for small P1′ residues such as glycine or alanine (or even serine or aspartate); this would be in agreement with the frequent occurrence of small P1′ residues in natural protein substrates (see, e.g., Lottenberg et al., 1983) and with thrombin's preference for alanine and glycine over leucine at this substrate position (Liem & Scheraga, 1974). P2′ side chains of protein inhibitors would normally interact with a surface provided by the side chain of residue 151 of the cognate serine proteinase (see Fig. 24). A proline residue is present at the P2′ position of the fibrinogen Aα-chain, and synthetic peptides with a proline at P2′ have been found to be readily cleaved by thrombin (Liem & Scheraga, 1974); in light of the presumed importance of the hydrogen bond formed between the P2′ amide nitrogen and the carbonyl group of residue 41 to achieve the transition state in productive enzyme–substrate complexes (Bode & Huber, 1986; Read & James, 1986), this preference for proline residues (which cannot mediate such a hydrogen bond) is remarkable. Inspection of the structure shows that this P2′ proline could be accommodated in a hydrophobic pocket (left to Leu 41 [Fig. 24]) with the following residues of a thrombin-bound peptide probably taking a different route than the binding loop of canonical protein inhibitors (Stubbs

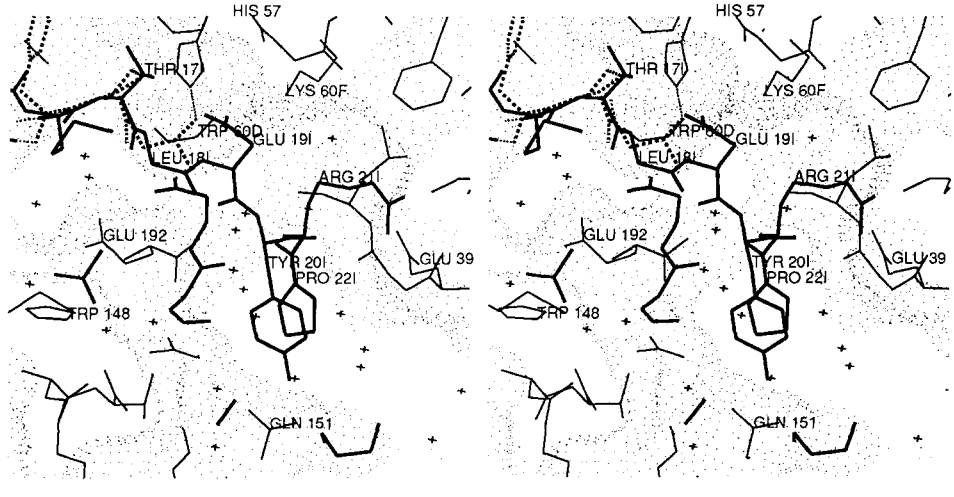

Fig. 24. Interface of the hypothetical complex of human thrombin with a Kazal-type inhibitor. The primed subsites of human thrombin (thin connections) are displayed together with the Connolly surface of thrombin; superimposed are the PPACK molecule (dashed connections) and part of the turkey ovomucoid third domain (thick connections) as determined in its complex with human leukocyte elastase (Bode et al., 1986) after fit of the elastase component to thrombin. The inhibitor scissile peptide bond is between Leu 181 (P1) and Glu 191 (P1′). The side chain of Glu 191 would collide with the indole ring of Trp 60D of the (partially cut-off) 60 insertion loop; only a shorter side chain should become accommodated at S1′. There is space for a Tyr 201 (P2′) side chain; the interactions with thrombin would, however, not be optimal. The side chain of Arg 211 (P3′) would collide with Glu 39; however, other more favorable arrangements can be assumed. The view is similar to the standard view of Figure 1.

et al., 1992). The relatively beneficial effect of an arginine residue at P3′ (Liem & Scheraga, 1974) could then be attributed to ionic interactions with glutamic acid residues 192 and 39 of thrombin (Fig. 24), in agreement with recent results obtained with a Glu 39 → Arg mutant (LeBonniec et al., 1991), and that of a valine residue at P4′ by hydrophobic interactions with Phe 60H.

Probable interaction with macromolecular substrates and inhibitors

The narrow, canyonlike structure of the thrombin active-site cleft is clearly the main source of the remarkable thrombin specificity toward macromolecular substrates and inhibitors (Bode et al., 1989b). It restricts access to the thrombin catalytic center of potential scissile peptide bonds of most macromolecules simply through steric hindrance, thereby discriminating in this way susceptible substrates by means of shape and conformation. An attempt to dock the archetypal inhibitor bovine pancreatic trypsin inhibitor (BPTI) to the thrombin active site with identical position and orientation as observed for the trypsin(ogen) complexes (Fig. 25) (Huber et al., 1974; Bode et al., 1978, 1984; Huber & Bode, 1978) results in a collision of residues of the inhibitor-binding region (in particular Tyr 35I and segment 37I–40I) with part of the 60 insertion loop of thrombin (in particular with the side chain of Trp 60D [see Fig. 25]).

It appears as if this steric hindrance cannot be relieved, even upon considerable relative tilting of the inhibitor. The inherently flexible Trp 148 loop (left-hand side in Fig. 25) could give way; the 60 insertion loop (right-hand side, Fig. 25), however, appears so rigid that strong forces would be required to allow access of the inhibitor. These docking experiments agree with experimental results, which reveal extremely low association constants of about $10^3\ M^{-1}$ for α-thrombin, but almost 10 times larger values for the more open and flexible γ-thrombin (Berliner et al., 1986; Ascenzi et al., 1988). The structure of the complexes formed at very high BPTI concentrations is unclear; however, an enforced expulsion of the 60 insertion loop by BPTI (in a similar manner to when the noncomplementary binding site of trypsinogen is forced into a fitting trypsinlike structure upon BPTI complexation [see Bode et al., 1978; Huber & Bode, 1978; Bode, 1979]) and consequent binding at the expense of free energy cannot be ruled out.

Similar docking experiments with other "small" serine proteinase protein inhibitors of known spatial structure such as eglin c (Bode et al., 1987), the squash seed inhibitor CMTI-I (Bode et al., 1989a), and the turkey ovomucoid third domain (Bode et al., 1986) show similar restraints (see Fig. 24). In each case the binding loop residues on both sides of the scissile peptide bond would essentially fit to the corresponding subsites around the thrombin catalytic site; however, collision with the thrombin "rims" would occur through structural elements extending from the inhibitor core. Of the protein inhibitors tested for regard to possible thrombin inhibition so far, the Kazal-type inhibitors seem to cause least steric hindrance (in spite of the noninhibiting behavior reported for ovomucoid [Berliner et al., 1986]); they therefore seem to be the most promising candidates for future protein design experiments.

In the absence of an experimental structure of an intact functional serpin inhibitor or of a serpin–serine proteinase complex, one can only speculate on the interaction mode between thrombin and antithrombin III. There are, however, several lines of evidence (see Bode & Huber, 1991) that serpins bind in an essentially similar manner as the "small" protein inhibitors to their cognate enzymes,

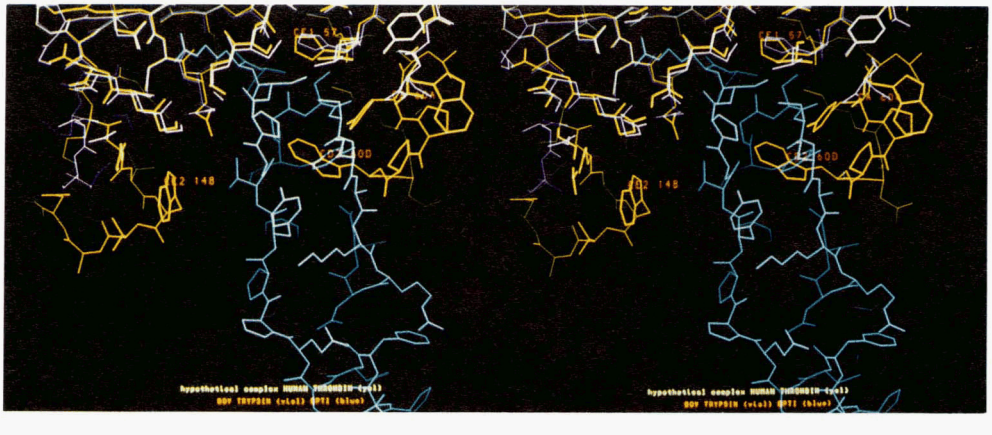

Fig. 25. Part of the hypothetical docking complex formed between human α-thrombin (yellow) and BPTI (blue). The BPTI docking is as found in the trypsin–Arg 15–BPTI complex (Bode et al., 1984); part of the trypsin component superimposed on thrombin is given in violet color. Only part of the cleft regions of thrombin and trypsin are shown; the full structure of BPTI is given; the view is along the active-site cleft of thrombin, which points downward. Clearly visible is the collision of Trp 60D with Tyr 35I of BPTI and the overcrowding, which would be expected in the surrounding region.

namely through an exposed binding loop of canonical conformation (Huber & Bode, 1978; Laskowski & Kato, 1980; Huber & Carrell, 1989); furthermore, the interaction of this loop with thrombin seems to be restricted essentially to the immediate environment of the active-site residues (Chang et al., 1979; Hofsteenge et al., 1988). Thrombin obviously imposes particular restrictions on the shape of the serpin-binding region, which must be flat. The core of these serpin inhibitors (Löbermann et al., 1984; Baumann et al., 1991) is much larger than those of the "small" inhibitors; therefore, more intimate molecular interactions with rim elements of the thrombin active-site canyon are to be expected.

Similar geometric requirements must be met by coagulation factors susceptible to thrombin action. The thrombin interaction with fibrinogen presumably does not follow this general interaction scheme, however. Recent crystallographic studies (Martin et al., 1992; Stubbs et al., 1992) reveal that segment Phe 8–Arg 16 of the fibrinopeptide A binds in a multiturnlike conformation to the hydrophobic cleft area of the thrombin active-site region (as suggested earlier by Scheraga and coworkers). With regard to the Aα-chain region carboxy-terminal of the scissile peptide bond Arg 16–Gly 17, one can only speculate that segment 30–44 (with several acidic residues) might bind to the fibrin(ogen) secondary binding exosite (Fig. 14) in a similar manner as seen for the bound hirudin tail (Hofsteenge et al., 1988).

Hirudin, however, interacts with thrombin in a quite different manner (Grütter et al., 1990; Rydel et al., 1990) to the canonically binding protein inhibitors (Fig. 26; see also Bode & Huber, 1991). It binds in a complementary manner with its large amino-terminal domain to thrombin-specific sites outside the enzyme's active site; only the (originally flexible [Folkers et al., 1989; Haruyama & Wuthrich, 1989]) amino-terminal segment comprising the first three hirudin residues becomes fixed under formation of a parallel β-pleated sheet (in contrast to PPACK [Fig. 21] and the canonical inhibitors). In addition, it associates with part of the active-site cleft and the positively charged fibrin(ogen) secondary binding exosite surface (Fig. 14) under formation of several surface-located salt bridges and hydrophobic contacts. Thus, the thrombin–hirudin interaction gains its tremendous specificity primarily through interactions with thrombin sites not common to other serine proteinases; due to the partial occlusion of substrate-binding sites and shielding of the active-site residues (not themselves directly blocked [Rydel et al., 1991]), both macromolecular and small substrates are denied access to the thrombin active site.

Materials and methods

Data collection and processing

Highly purified human α-thrombin was a kind gift of Drs. Jan Hofsteenge and S.R. Stone, Friedrich-Miescher-Institut, Basel. PPACK was purchased from Calbiochem or Bachem. Formation of the PPACK–human α-thrombin complex and crystallization in polyethylene glycol 6000/phosphate buffer, pH 7, were described previously (Bode et al., 1989b). The crystals belong to the orthorhombic space group $P2_12_12_1$ and contain one molecule per asymmetric unit with cell constants a = 87.74 Å, b = 67.81 Å, and c = 61.07 Å. Larger crystals initially show reflections to a minimal Bragg spacing of beyond 1.9 Å resolution. The crystals were transferred to a mother liquor of 30% polyethylene glycol 6000, buffered with 0.2 M sodium/potassium phosphate to pH 7.0 for data collection.

At first, a particularly large, but macroscopically twinned crystal obtained by microdialysis was cut into four parts; two of them were used initially for data collection on a rotation camera (see below). Only one of these crystal fragments (of size 0.20 × 0.4 × 0.6 mm^3) yielded completely unsplit reflections (initially to 1.9 Å) and was transferred to a FAST television area detector diffractometer (Enraf-Nonius) using Ni-filtered CuKα radiation, a crystal-to-detector distance of 40 mm, and a detector angle of 10°. An X-ray data set to 2.45 Å resolution (Table 12) was collected upon rotation of the crystal fragment of 92° about its b axis and of 30° about an axis inclined by 30° to b. Capillary and crystal were kept in a cold air stream at 3 °C. The program package

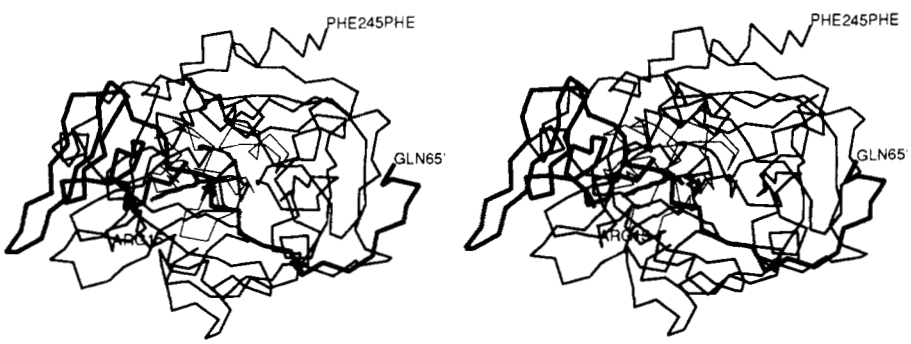

Fig. 26. α-carbon structures of hirudin (bold connections) and thrombin (mediate connections) (Rydel et al., 1991), superimposed with BPTI (thin connections) as it binds to trypsin in the trypsin–Arg 15–BPTI complex (Bode et al., 1984). The P1 residue of Arg 15–BPTI, Arg 15I, is shown with dashed connections.

Table 12. *Intensity data statistics*

A. Fast data		
Number of evaluated reflections		40,583
*R*merge[a]		0.109
Number of Friedel mates		18,684
*R*merge[a] of Friedel mates		0.047
Number of independent reflections		10,797
Maximal resolution		2.45 Å
Measured/expected reflections	316 to 2.45 Å	0.80
B. Photographic films		
Number of evaluated reflections		26,009
*R*merge[a]		0.071
Number of independent reflections		14.105
Maximal resolution		1.92 Å
C. Combined data		
Number of evaluated reflections		66,592
Number of independent reflections		16,910
Maximal resolution		1.92 Å
*R*merge after rejection[b]		0.101
Completeness		
Measured/expected reflections	316 to 2.45 Å	0.87
	2.55 to 2.45 Å	0.70
Measured/expected reflections	316 to 2.2 Å	0.77
	2.25 to 2.2 Å	0.42
Measured/expected reflections	316 to 1.92 Å	0.65
	2.05 to 1.92 Å	0.16

[a] $R\text{merge} = \sum h \sum i[I(i,h) - \langle I(h)\rangle]/\sum h\langle I(h)\rangle$, where $I(i,h)$ is the intensity of reflection h observed for the ith source, and $\langle I(h)\rangle$ is the mean intensity of reflection h for all measurements of $I(h)$.

[b] Individual measurements were rejected if their relative deviation from the mean value $(I - \langle I\rangle/\langle I\rangle)$ exceeded the threshold limit of 0.6.

MADNES (Messerschmidt & Pflugrath, 1987) was used to determine and refine the crystal setting parameters, crystal cell constants, and camera parameters and to evaluate the reflection data on-line. All intensities with three contiguous pixels greater than 3.5σ above background were assigned as observed intensities. These data were corrected for absorption effects using equivalent reflections and a nonuniformity table based on previously collected data to determine the absorption ellipsoid of the crystal (Messerschmidt et al., 1990). The Friedel-mates (Table 12) were loaded, scaled, and merged using the program system PROTEIN (Steigemann, 1974). Number, quality, and completeness of these FAST data are shown in Table 12.

Additional high-resolution X-ray data up to 1.92 Å resolution (Table 12) were collected on a rotation camera (Huber, Rimsting) in the cold room at 4 °C using graphite monochromatized CuKα radiation. For data collection the large crystal later used for FAST data collection (see above), another slightly disordered "twin"-brother, and three additional, smaller crystals (of approximate size $0.05 \times 0.2 \times 0.5$ mm^3) were rotated about 15° each around their b axes. In this way the whole reciprocal space was systematically surveyed. Due to radiation damage the maximal resolution of the larger crystals decreased to about 2.2 Å upon X-ray exposure according to rotation photographs.

The films were scanned on an Optronics rotating-drum film scanner and evaluated with the program FILME as modified by W.S. Bennett, Jr. Reflections with intensity $>1.0\sigma$ were further processed. Some of the film data obtained from the smaller vapor diffusion-made crystals fitted poorly to the FAST data; in spite of individual *R*sym values as low as 0.06, some of them exhibited *R*merge values (defined as $\sum(|I - \langle I\rangle|)/\sum|I|$) above 0.20. Film data from one of these crystals and from all films with *R*merge values >0.20 were therefore omitted from the final merged data set. Table 12 gives only the intensity statistics for accepted camera data, which represent about 70% of the asymmetric unit of reciprocal space.

The combined and scaled data set comprises 16,910 independent X-ray data. Completeness is almost full down to 2.45 Å and gradually decreases with increasing resolution to 1.92 Å (Table 12).

Search for orientation and position

Figure 27 gives a schematic arrangement of the final thrombin model (see also Fig. 3). Highlighted are those amino acid residues that were selected to form the truncated model used for the search in the PPACK-thrombin crystals. Due to the anticipated closest topological similarity with thrombin, the polypeptide backbone of the bovine chymotrypsin model (Brookhaven Data Bank [Bernstein et al., 1977] codes 2CGA [Cohen et al., 1981], 4CHA [Tsukada & Blow, 1985], 5CHA [Blevins & Tulinsky, 1985]) was used as a guiding structure. Some structural features of bovine trypsin (Bode & Schwager, 1975a; Bode et al., 1976) expected to be closer at a few sites were incorporated into the truncated model as well. Ninety-seven amino acid residues of human α-thrombin (emphasized by filled circles in Fig. 27) were assumed to be identical or almost identical to residues in chymotrypsin and were used in their entirety; another 105 residues (indicated by bold circles) were considered to be topologically equivalent to either chymotrypsin or trypsin and were included only with their main-chain and side-chain atoms up to Cβ. No efforts were made to construct a contiguous energy-minimized search model.

In addition, the truncated model contained also the two last residues of PPACK including the hemiketal group (modeled according to the peptidyl chloromethylketone moiety of inhibited human leukocyte elastase [Wei et al., 1988]) and 10 conserved internal solvent molecules. The truncated model used for subsequent searches consisted of 1,352 active atoms, corresponding to slightly more than half of all atoms of the final peptidyl structure. This residue selection corresponds approximately to the structurally conserved regions (SCR, CR) defined by Greer (1981) and Bing et al. (1986).

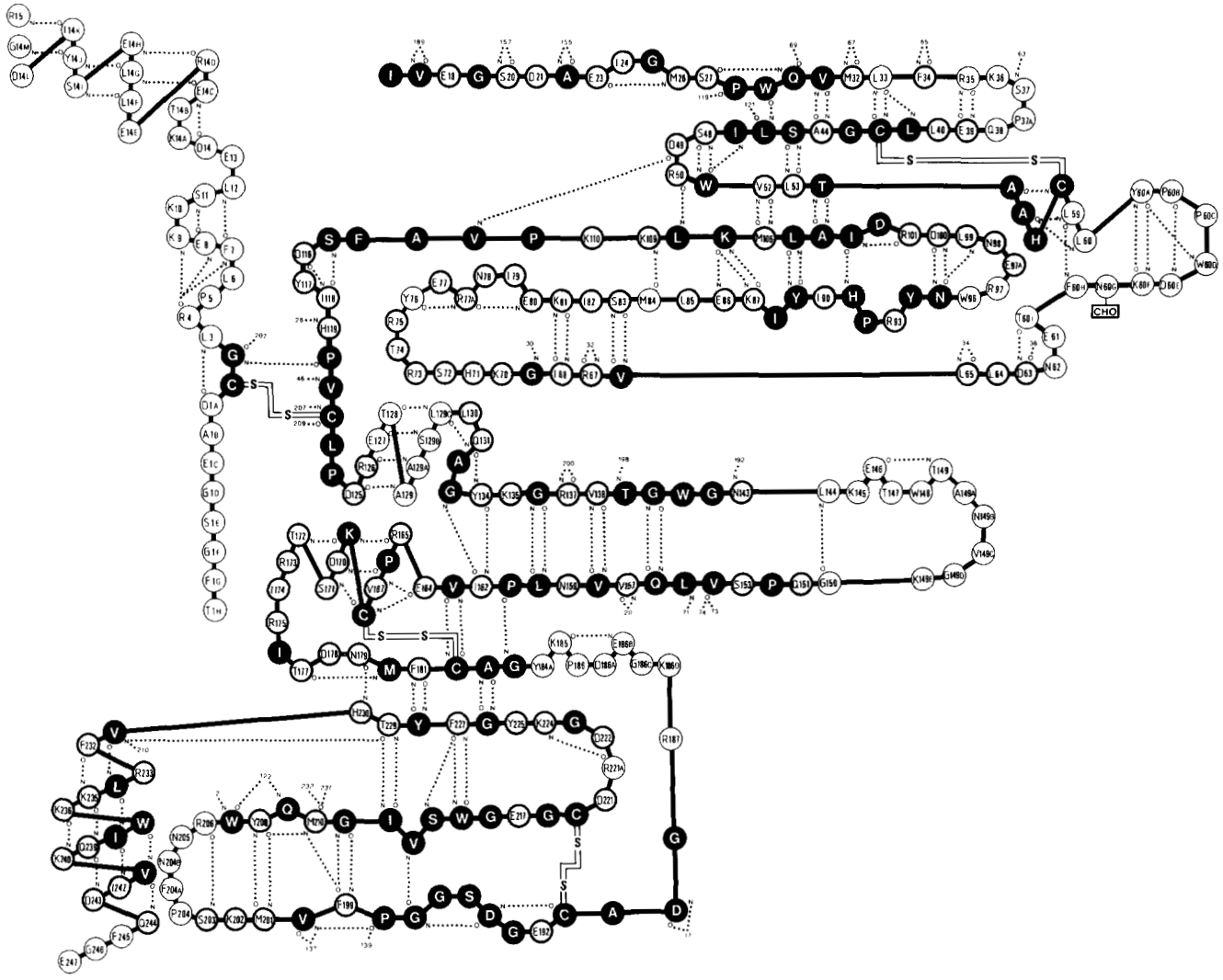

Fig. 27. Schematic representation of the truncated model used for search; the thrombin chain segments are arranged as found for the final thrombin molecule (see Fig. 3). The backbone atoms plus Cβ of 202 residues of bovine chymotrypsin (in a few cases: of bovine trypsin, see text) assumed to be topologically equivalent to thrombin (emphasized by thick circles) were included; of these 202 residues, 97 given by filled circles were taken with full size.

The orientation of the thrombin B-chain in the thrombin crystals was determined by Patterson search methods. The truncated model was placed in a triclinic cell with orthogonal cell axes of length 120 Å. A model Patterson map was calculated with data between 8 and 3.5 Å using a B value of 20 Å^2; grid points representing the 2,463 highest function values within a radial shell of 3 and 15 Å were selected. A crystal Patterson synthesis was calculated using reflections between 8 and 3.5 Å resolution on a 1-Å grid. The rotational search was carried out by calculating the product correlation function of the crystal and the modified model Patterson map as a function of the orientational angles θ_1, θ_2, θ_3 (defined as subsequent rotations around z, the new x [x'] and the new z [z'' axis]) in 5° steps in the angle range θ_1 = 0–90°, θ_2 = 0–180°, θ_3 = 0–180° (with the PROTEIN program system of Steigemann [1974]). The rotation function showed a prominent peak of height 101.2 arbitrary units (corresponding to 5.6σ above the mean value of 70.5) at 10°/125°/90° with second and third peak heights of 92.7 and 92.6 (4.1 and 4.1σ), respectively. A fine scan in 0.2° intervals increased the value of the highest peak to 103.2 units, yielding an optimal orientation value.

The location of the thrombin B-chain with respect to the crystallographic symmetry elements was determined using the translation function and programs written by E.E. Lattmann and modified by J. Deisenhofer and R. Huber. An 8- to 3-Å model transform was calculated for the truncated search model oriented according to the rotation function solution and placed with its center of gravity at the origin of a triclinic cell with orthogonal cell axes of length 120 Å. Only one of the three cross-vector

sets yielded the highest peak on a Harker section; all three highest Harker section peaks yielded, however, a consistent solution (Table 13). The crystallographic R-value (defined as $\sum(|F\text{obs} - F\text{calc}|)/\sum|F\text{obs}|$) of the appropriately oriented and positioned truncated model was 0.53 for 8- to 3-Å data.

The orientation and position parameters of this model were further refined with the Fourier transform fitting program TRAREF (Huber & Schneider, 1985) resulting in rotational and positional shifts of about 3° and 0.25 Å. The R-value of this model was 0.493 for 8- to 3-Å data.

Model building and crystallographic refinement

A $2F\text{obs} - F\text{calc}$ Fourier map was calculated with this model by means of PROTEIN using intensity data from 8 to 3 Å and Sim-weighted model phases. This map displayed most parts of the core-forming segments and of the PPACK moiety; in addition it showed electron density for several omitted side chains (such as Phe 1I), for some insertion loops (such as loop 202–208), and for a short A-chain segment following Cys 1–Gly 2. These defined structural parts were modeled into density using the PSFRODO version (Pflugrath et al., 1984) of FRODO (Jones, 1978) on a PS390 display system (Evans & Sutherland). This new, more complete model had a decreased R-value of 0.44.

This model was subjected to 12 macrocycles of modeling and cycles of crystallographic refinement (Table 14). Most refinement was carried out using the energy-restrained least-squares program EREF (Jack & Levitt, 1978). Each EREF macrocycle started with inspection of Fourier and difference Fourier (and sometimes "omit") maps at the display with appropriate model manipulations and additions, followed by 5–9 refinement cycles, each consisting of calculation of structure factors and of a Sim-weighted difference Fourier (with PROTEIN), calculation of the normal equations from the difference Fourier (DERIV [Jack & Levitt, 1978]), and subsequent least-squares refinement of positional parameters with energy restraints (EREF [Jack & Levitt, 1978]). In between, thermal parameters were also refined. Most standard geometries were taken from Levitt (1974); a few improper angles were introduced to keep carboxylate and carboxamide groups planar; some peptide bond parameters were updated, and the proline parameters were changed to allow ring puckering (see Schirmer et al., 1987).

The carboxy-terminal arginine methyl residue of PPACK (Arg 3I) was constructed from an EREF stan-

Table 13. *Position and height (in arbitrary units) of highest peaks in the translation function*

Vector set	Peak height	Standard deviation	Peak position $u/80$	$v/70$	$w/60$	Harker section
1 → 2	382.1		40.3	55.1	30.0	
	374.5	58.4	8.3	2.3	0.0	$w = 30$
	373.9		9.6	50.3	0.0	
1 → 3	683.1		0.0	4.4	5.1	
	664.4	102.3	40.0	20.1	50.8	$u = 40$
	561.5		0.0	10.9	51.1	
1 → 4	536.6		5.7	0.0	3.5	
	526.6	78.8	0.0	35.0	20.3	$v = 35$
	430.5		5.3	0.0	59.0	

Table 14. *Course of crystallographic refinement*

Macrocycle	Active atoms	Refinement method	Resolution (Å)	Internal energy (kcal/mol)	r.m.s. bond deviation from target values	R-value
						0.493
0	1,352	–	8 to 3	–	–	0.444
1	1,531	EREF/DERIV	8 to 3	−488	0.014	0.345
2	1,765	EREF/DERIV	8 to 2.5	−320	0.014	0.329
3	1,889	EREF/DERIV	8 to 2.5	−291	0.015	0.289
4	2,070	EREF/DERIV	6 to 2.5	−420	0.014	0.251
5	2,205	EREF/DERIV	6 to 2.5	−472	0.016	0.234
6	2,331	EREF/DERIV	6 to 2.2	−752	0.016	0.234
7	2,365	EREF/DERIV	6 to 1.92	−1,056	0.018	0.200
8	2,391	EREF/DERIV	6 to 1.92	−1,362	0.018	0.178
9	2,589	EREF/DERIV	6 to 1.92	−1,298	0.018	0.173
10	2,650	EREF/DERIV	6 to 1.92	−1,227	0.019	0.171
11	2,833	X-PLOR	8 to 1.92	−2,857	0.013	0.174
12a	2,819	X-PLOR	8 to 1.92	−3,180	0.013	0.156
12b	2,833	EREF	8 to 1.92	−1,442	0.018	0.159

dard arginine residue by adding a methylene group and enforcing a tetrahedral geometry at the carbonyl carbon as done for the complex of human leukocyte elastase with a peptidyl chloromethylketone (Wei et al., 1988). All nonbonded restraint parameters for Ser 195 Oγ and for the Arg 3I methylene group were set to zero to allow their unbiased approach to the Arg 3I carbonyl group and to the His 57 Nε2 atom, respectively. At later stages of refinement, solvent molecules interpreted as water molecules were inserted at stereochemically reasonable positions if the difference electron density exceeded 4σ (using PROTEIN routines).

The graphic program MAIN (D. Turk) was employed in macrocycle 11 in combination with the conjugate gradient minimization procedure of X-PLOR version 1.5 (Brünger et al., 1987). Crystallographic refinement with X-PLOR involved 50 cycles of positional refinement followed by 50 steps of B-value optimization. New solvent molecules were searched for by means of MAIN routines that checked the difference Fourier map for peaks greater than 2.5σ and inserted waters (closest to the original molecule) if located within hydrogen-bond distance (1.8–3.5 Å) of polar protein atoms or of other solvent molecules; current solvent molecules were checked for height of density in the Fourier map and were deleted if positioned in density less than 0.8σ above the mean value. The X-PLOR forcefield parameters given in PARAM19X.PRO (which treats only polar hydrogens explicitly) were used for restrained crystallographic refinement; all side-chain charges were set to zero (as recommended in the X-PLOR manual), and all histidine residues were treated as singly protonated, with His 57 carrying a proton at Nδ1 and all other histidines at Nε2.

PPACK residue Arg 3I was modeled according to an X-PLOR standard arginine residue; additional patch residues were built to allow covalent linkages to the preceding residue (Pro 2I), via the carbonyl carbon (CF) to Ser 195 Oγ and via its methylene carbon (CJ) to His 57 Nε2, respectively. The hemiketal group of Arg 3I was restrained to an ideal tetrahedral state (with all angles around the carbonyl carbon [CF] being 109.5°); the lengths of both CF-O bonds (from the carbonyl carbon to its oxygen and to Ser 195 Oγ) were restrained to 1.42 Å (as C-O single bonds), and the bond length between the methylene carbon and His 57 Nε2 to 1.49 Å (taken from X-PLOR parameters).

The course of the crystallographic refinement is shown in Table 14. The density calculated with the first, already improved model displayed spurious density for most of the insertion loops, which only from macrocycle 3 on became clear enough to be modeled. From macrocycle 4 on, the full B-chain was present; several peptide segments, in particular some insertion loops and the carboxy-terminus, remained subject to changes until the very end of refinement. From macrocycle 6 on, additional water molecules were localized and included in refinement. At macrocycle 8 the A-chain was clearly traceable from Asp 1A to Tyr 14J; the preceding residues Ala 1B and Glu 1C were modeled and refined with relatively bad conformations; handicapped by several breaks in density, the first five residues, Thr 1H to Gly 1D, were tentatively placed along a shallow groove along the B-chain surface. At this stage no continuous density was visible for the last four residues, Ile 14K–Arg 15, of the A-chain.

The model resulting from macrocycle 10 was that described recently (Bode et al., 1989b). It comprised 290 defined residues (lacking the four carboxy-terminal residues of the A-chain and the carboxy-terminal Glu 247 of the B-chain) with some side chains remaining partially undefined, and 291 solvent molecules, and had an *R*-value of 0.171 (6–1.92-Å data).

Upon further inspection of the maps it became clear that some hitherto undefined side chains could be localized and furthermore that Lys 110 (with a very mobile side chain) had to be rotated and that the density "tongue" extending away from the main chain and being formerly interpreted as the Lys 110 side chain should rather accommodate its carbonyl group; as a consequence, the peptide bond formed with the succeeding prolyl residue, Pro 111, should (contrary to our previous suggestion) be in a trans rather than in a cis conformation. Thus, only one of the proline residues of human thrombin (Pro 37) is in a cis conformation, in full agreement with the structures of the thrombin–hirudin complex (Rydel et al., 1991). In addition, a slightly different pathway for the A-chain segment preceding Glu 1C was found; as a result, the peptide segment from Thr 1H to Glu 1C is still arranged in a multiple-turn conformation along the same B-chain surface as before, but is now slightly better defined by density with a more reasonable stereochemistry. Likewise, some faint but (at very low contouring) continuous density allowed the placement of the carboxy-terminal four A-chain residues and Glu 247 of the B-chain with more confidence and to include all side-chain atoms assigned before as dummy atoms in phase calculation.

The *R*-value of this partially revised thrombin model was 0.193. A subsequent restrained refinement with X-PLOR (macrocycle 11 [Table 14]) using the default bond and angle B-value restraints brought it down to 0.174. In a final unrestrained B-value refinement with X-PLOR (macrocycle 12a) or EREF (macrocycle 12b) this model converged (without rejection of any bad reflections) to final *R*-values of 0.156 and 0.159, respectively. The final parameters of both models obtained from X-PLOR and EREF refinement are given in Table 15. The models are virtually identical (with an r.m.s. deviation of 0.07 Å between all 295 α-carbon atoms, i.e., well below the estimated accuracy of each model); the small differences result from different target and restraint parameters in both procedures.

Table 15. *Final model parameters of PPACK-thrombin*

	EREF		X-PLOR
Number of active thrombin atoms		2,380	
Number of active PPACK atoms		30	
Number of active solvent atoms	423		409
r.m.s. standard deviation from target values			
Bond lengths	0.018 Å		0.013 Å
Bond angles	3.085°		3.137°
Total energy	−1,442 kcal/mol		−3,733 kcal/mol
Resolution range		8 to 1.92 Å	
Number of reflections used in refinement		15,019	
R value	0.159		0.156
Mean B value			
Whole data set	26.8 Å^2		29.9 Å^2
Thrombin nonhydrogen atoms	22.9 Å^2		26.8 Å^2
PPACK nonhydrogen atoms	8.4 Å^2		10.2 Å^2
Solvent molecules	45.3 Å^2		49.7 Å^2
Estimated mean coordinate error from agreement between observed and calculated structure factor amplitudes			
According to Luzzati (1952)	0.23 Å		0.22 Å
According to Cruickshank (1949)	0.11 Å		0.11 Å

The B-values for each thrombin residue, averaged over all nonhydrogen atoms, are shown in Figure 28A; the mean electron density per residue is given in Figure 28B; with the exception of peptide bond 14L–14M (<0.6σ) and 1H–1G and 1F–1E (<0.8σ) the main chain is everywhere represented by continuous electron density above 1σ. In general, the oscillations reflect the alternate arrangement on the surface and in the interior of the molecule; particularly high B-values are observed for both A-chain termini, for the exposed B-chain loops 70–80 and 204A–205, and for the carboxy-terminus of the B-chain.

A very few noninterpreted maxima and minima remain in the solvent region of the final difference Fourier map with heights up to 5σ and −3σ, respectively; no new solvent molecules could be placed at stereochemically reasonable positions, however. According to the Ramachandran plot (Fig. 29) almost all main-chain angles of the thrombin model are within the allowed regions defined by Ramakrishnan and Ramachandran (1965). The angle pairs outside belong to the weakly defined amino- and carboxy-termini of the thrombin A-chain; they are labeled in Figure 29 and presumably reflect multiple conformations of the respective residues.

For such energy-restrained least-squares procedures it

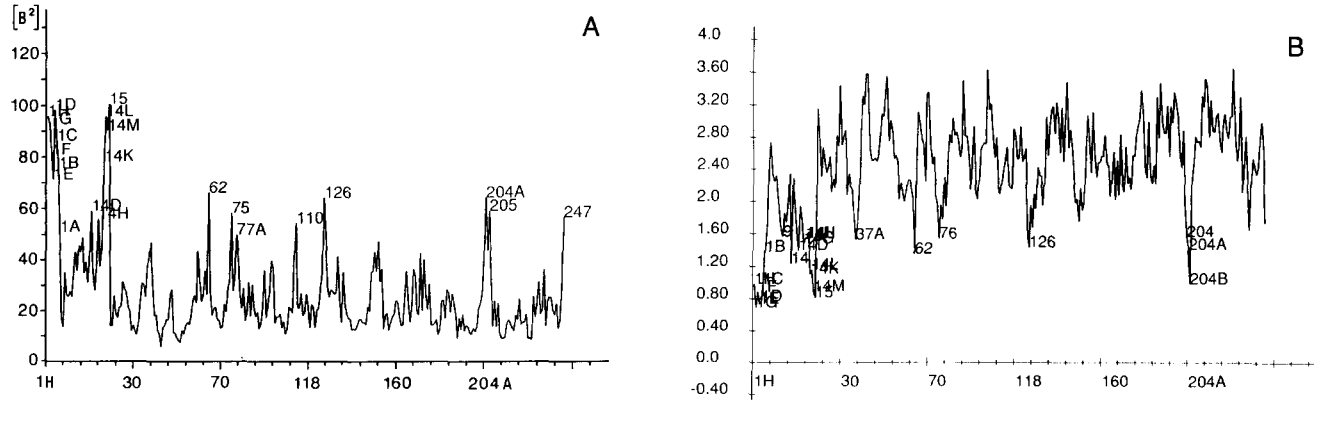

Fig. 28. **A**: Temperature factors of the final thrombin model averaged over all nonhydrogen atoms of each residue. **B**: Mean density per residue (in σ) of the final $2F$obs − Fcalc electron density map averaged over the main-chain atoms of each residue.

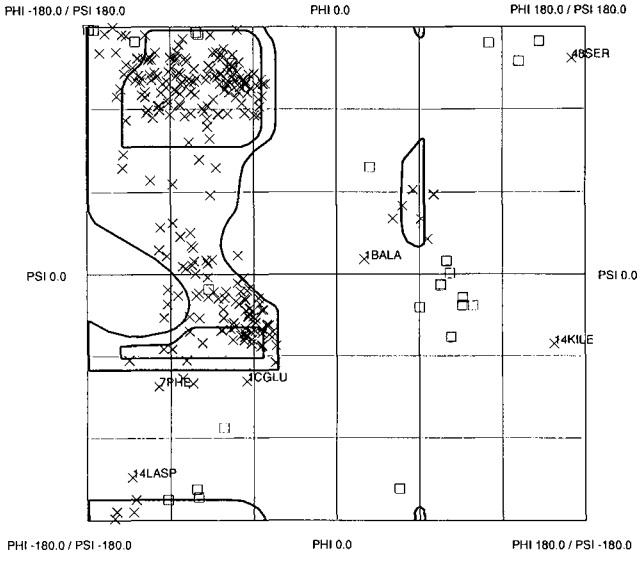

Fig. 29. Ramachandran plot of the refined thrombin model. O, glycyl residues; X, all other residues.

is difficult to define an adequate standard deviation for positional parameters of the atoms. From the scattering angle dependence of the *R*-values (Luzzati, 1952) the upper limit of the mean positional error of the refined model is estimated to about 0.21 Å (Fig. 30; Table 15). The formula derived by Cruickshank (1949) gives a lower limit of 0.11 Å for the accuracy of the final structure.

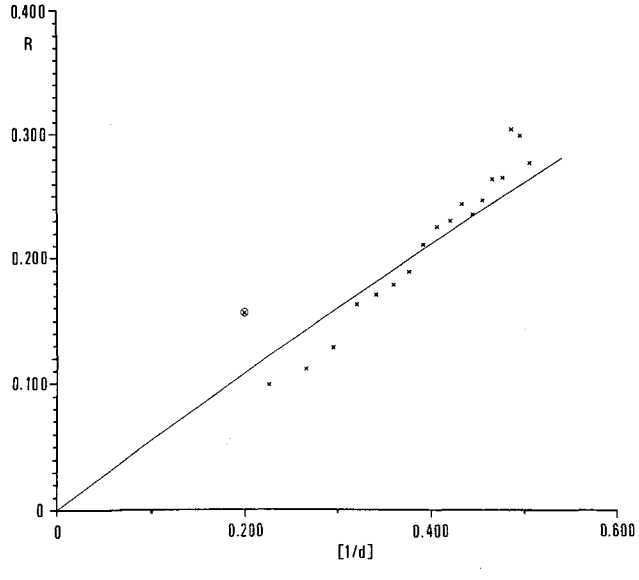

Fig. 30. Luzzati plot of the final thrombin model after X-PLOR refinement.

Hydrogen-bond definition, alignment procedures, and electrostatic calculations

Accessible surface calculations and identification of possible hydrogen bonds were carried out by means of the program HYBAC (Levitt, 1974). Possible hydrogen-bond donor–acceptor interactions were considered as hydrogen bonds if the interaction energy was $E < -0.7$ kcal/mol calculated with DSSP according to Kabsch and Sander (1983). Interface areas were calculated with a program of T.J. Richmond (see Richards, 1977). Most of the plots were made using the graphics facilities of FRODO, MAIN (D. Turk), and a RIBBON version (Priestle, 1988) modified by one of us (A.K.)

Optimal superposition of related proteinases and determination of the number of topologically equivalent α-carbon atoms was achieved with the program OVRLAP (Rossmann & Argos, 1975) as modified by W.S. Bennett; the following selection parameters were used: *P*cut = 0.1; $E_1 = E_2 = 2.0$; run length >6.

The amino acid sequences were aligned using the program system ALIGN (Needleman & Wunsch, 1970). In single runs only pairs of sequences were compared; the Mutation Data Matrix (250 PAMS) was used as scoring matrix; a penalty of 20 was assigned to each break, and 100 random runs were performed for each comparison. Protein sequence searches were done against the MIPSX merged Protein Sequence Data Base, Release 25.0, June 30, 1990 (Martinsried Institute for Protein Sequences).

The electrostatic calculations were performed on the basis of the modified Tanford-Kirkwood theory (Tanford & Kirkwood, 1957; Karshikov et al., 1989) as well as of the finite-difference algorithm developed by Gilson et al. (1985) and Klapper et al. (1986). In both methods the charges carried by the titratable groups and the peptide dipoles are considered. The point charges were attributed (1) to those between NH1 and 2 of the arginine guanidyl groups, (2) to Nζ of lysine residues, (3) to the center of imidazole rings, and (4) to those between both carboxylate oxygen atoms of aspartates and glutamates; orientation and size of the peptide dipoles were taken as described by Karshikov et al. (1989). The solvent accessibilities of the charged groups in the protein structures defined as the accessibility ratio to that of a fully extended Ala-Xaa-Ala peptide were calculated according to Lee and Richards (1981). Dielectric constants of 4 and 78 were used for the space inside and outside of the protein. All calculations were performed for an ionic strength of 0.125 and a pH value of 7.0. In the case of the Tanford–Kirkwood model the intrinsic pK values were taken from Matthew (1985).

Both methods yielded quite consistent results regarding the shape of the electric potential; the contouring surfaces shown in Figure 8 and the electrostatic interaction energies presented in this paper are those derived by

means of the Tanford-Kirkwood model. In the case of the finite-difference method, the calculations were started with coulombic boundary conditions followed by two steps of focusing calculations. The electrostatic equipotential surfaces were programmed (A.K.) to be displayed on a PS system using the FRODO graphics facilities.

Acknowledgments

We thank Prof. Dr. Robert Huber in particular for his long-standing interest in this work, Drs. J. Hofsteenge and S.R. Stone, Friedrich-Miescher-Institut, for human thrombin samples and for stimulating discussions, Dr. M. Stubbs for careful reviewing of this manuscript, and Drs. P. Ruecknagel, E. Jäger, and H. Oschkinat for preparation of and for NMR measurements with the YPPWDN peptide. The help of I. Mayr, U. Baumann, and other colleagues of the Abteilung für Strukturforschung and of the Martinsried computer center is acknowledged. The financial support by the Sonderforschungsbereich 207 (H1) of the University of Munich, by the Fonds der Chemischen Industrie, by the Alexander-von-Humboldt-Stiftung, and by the Research Council of Slovenia/Ministry for Science and Technology is gratefully acknowledged.

References

Ascenzi, P., Coletta, M., Amiconi, G., de Cristofaro, R., Bolognesi, M., Guarneri, M., & Menegatti, E. (1988). Binding of the bovine basic pancreatic trypsin inhibitor (Kunitz) to human α-, β- and γ-thrombin; a kinetic and thermodynamic study. *Biochim. Biophys. Acta 956*, 156–161.

Bajusz, S., Barabas, E., Tolnay, P., Szell, E., & Bagdy, D. (1978). Inhibition of thrombin and trypsin by tripeptide aldehydes. *Int. J. Peptide Protein Res. 123*, 217–221.

Bajusz, S., Szell, E., Bagdy, D., Barabas, E., Horvath, G., Dioszegi, M., Fittler, Z., Szabo, G., Juhasz, A., Tomori, E., & Szilagy, E. (1990). Highly active and selective anticoagulants: D-Phe-Pro-Arg-H, a free tripeptide aldehyde prone to spontaneous inactivation, and its stable *N*-methyl derivative, D-Me-Phe-Pro-Arg-H. *J. Mol. Chem. 33*, 1729–1735.

Bar-Shavit, R., Eldor, A., & Vlodavsky, I. (1989). Binding of thrombin to subendothelial extracellular matrix; protection and expression of functional properties. *J. Clin. Invest. 84*, 1096–1104.

Bar-Shavit, R., Kahn, A., Mann, K.G., & Wilner, G.D. (1986). Growth-promoting effects of esterolytically inactive thrombin on macrophages. *Cell Biochem. 32*, 261–272.

Bar-Shavit, R., Kahn, A., Mudd, M.S., Wilner, G.D., Mann, K.G., & Fenton, J.W., II (1984). Localization of a chemotactic domain in human thrombin. *Biochemistry 23*, 397–400.

Bar-Shavit, R., Kahn, A., Wilner, G.D., & Fenton, J.W., II (1983). Monocyte chemotaxis: Stimulation by a specific exosite region in thrombin. *Science 220*, 728–731.

Bar-Shavit, R., Sabbah, V., Lampugnani, M.G., Marchisio, P.C., Fenton, J.W., II, Vlodavsky, I., & Dejana, E. (1991). An Arg-Gly-Asp sequence within thrombin promotes endothelial cell adhesion. *J. Cell Biol. 112*, 335–344.

Baumann, U., Huber, R., Bode, W., Grosse, D., Lesjak, M., & Laurell, C.B. (1991). Crystal structure of cleaved human α_1-antichymotrypsin at 2.7 Å resolution and its comparison with other serpins. *J. Mol. Biol. 218*, 595–606.

Berg, W., Hillvärn, B., Arwin, H., Stenberg, M., & Lundström, I. (1979). The isoelectric point of thrombin and its behaviour compared to prothrombin at some solid surfaces. *Thromb. Haemostasis 42*, 972–982.

Berliner, L.J., Birktoft, J.J., Miller, T.L., Musci, G., Scheffler, J.W., Shen, Y.Y., & Sugawara, Y. (1986). Thrombin: Active-site topography. *Ann. N.Y. Acad. Sci. 485*, 80–95.

Berliner, L.J. & Shen, Y.Y.L. (1977). Physical evidence for an apolar binding site near the catalytic center of human α-thrombin. *Biochemistry 16*, 4622–4626.

Berliner, L.J., Sugawara, Y., & Fenton, J.W., II (1985). Human α-thrombin binding to nonpolymerized fibrin-sepharose: Evidence for an anionic binding region. *Biochemistry 24*, 7005–7009.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rogers, J.R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The protein data bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol. 122*, 535–542.

Bezeaud, A. & Guillin, M.C. (1988). Enzymic and nonenzymic properties of human β-thrombin. *J. Biol. Chem. 263*, 3576–3581.

Bing, D.H., Cory, M., & Fenton, J.W., II (1977). Exosite affinity labeling of human thrombins; similar labeling on the A chain and B chain fragments of clotting α- and non-clotting γ/β thrombins. *J. Biol. Chem. 252*, 8027–8034.

Bing, D.H., Feldmann, R.J., & Fenton, J.W., II (1986). Structure and function relationships of thrombin based on the computer generated three-dimensional model of the B chain of bovine thrombin. *Ann. N.Y. Acad. Sci. 485*, 104–119.

Bizios, R., Lai, L., Fenton, J.W., II, & Malik, A.B. (1986). Thrombin-induced chemotaxis and aggregation of neutrophils. *J. Cell. Physiol. 128*, 485–490.

Björk, I. & Lindahl, V. (1982). Mechanism of the anticoagulant action of heparin. *Mol. Cell. Biochem. 48*, 161–182.

Blevins, R.A. & Tulinsky, A. (1985). The refinement and the structure of the dimer of α-chymotrypsin at 1.67 Å resolution. *J. Biol. Chem. 260*, 4264–4275.

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

Blombäck, B., Hessel, H., Hogg, D., & Claesson, G. (1977). Substrate specificity of thrombin on proteins and synthetic substrates. In *Chemistry and Biology of Thrombin* (Lundblad, R.L., Fenton, J.W., II, & Mann, K.G., Eds.), pp. 275–290. Ann Arbor Science Publishers, Ann Arbor, Michigan.

Bode, W. (1979). The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. II. The binding of the pancreatic trypsin inhibitor and of isoleucine-valine and of sequentially related peptides to trypsinogen and to *p*-guanidinobenzoate-trypsinogen. *J. Mol. Biol. 127*, 357–374.

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 proteinase. Crystallization, structure determination, crystallographic refinement, structure and its comparison with bovine trypsin. *J. Mol. Biol. 164*, 237–282.

Bode, W., Greyling, H.J., Huber, R., Otlewski, J., & Wilusz, T. (1989a). The refined 2.0 Å X-ray crystal structure of the complex formed between bovine β-trypsin and CMTI-I, a trypsin inhibitor from squash seeds (*Cucurbita maxima*). Topological similarity of the squash inhibitors with the carboxypeptidase A inhibitor from potatoes. *FEBS Lett. 242*, 285–292.

Bode, W. & Huber, R. (1986). Crystal structures of pancreatic serine endopeptidases. In *Molecular and Cellular Basis of Digestion* (Desnuelle, P., Ed.), pp. 213–234. Elsevier, Amsterdam, New York, Oxford.

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

Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R., & Hofsteenge, J. (1989b). The refined 1.9 Å crystal structure of human α-thrombin: Interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J. 8*, 3467–3475.

Bode, W., Papamokos, E., & Musil, D. (1987). The high resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and the inhibitor eglin C, an elastase inhibitor from the leech *Hirudo medicinalis*. Structural analysis, subtilisin structure and interface geometry. *Eur. J. Biochem. 166*, 673–692.

Bode, W. & Schwager, P. (1975a). The refined crystal structure of bovine β-trypsin at 1.8 Å resolution. II. Crystallographic refinement,

calcium binding site, benzamidine binding site and active site at pH 7. *J. Mol. Biol. 98*, 693–717.
Bode, W. & Schwager, P. (1975b). The single calcium binding site of crystalline bovine β-trypsin. *FEBS Lett. 56*, 139–143.
Bode, W., Schwager, P., & Huber, R. (1976). Structural studies on the pancreatic trypsin inhibitor–trypsin complex and its free components: Structure and function relationship in serine protease inhibition and catalysis. In *Proteolysis and Physiological Regulation*, pp. 43–76. Miami Winter Symposia. Academic Press, New York, San Francisco, London.
Bode, W., Schwager, P., & Huber, R. (1978). The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. The refined crystal structures of bovine trypsinogen–pancreatic inhibitor complex and of its ternary complex with Ile-Val at 1.9 Å resolution. *J. Mol. Biol. 118*, 99–112.
Bode, W., Turk, D., & Stürzebecher, J. (1990). Geometry of binding of the benzamidine- and arginine-based inhibitors NAPAP and MQPA to human α-thrombin. X-ray crystallographic determination of the NAPAP–trypsin complex and modeling of NAPAP–thrombin and MQPA–thrombin. *Eur. J. Biochem. 193*, 175–182.
Bode, W., Walter, J., Huber, R., Wenzel, H.R., & Tschesche, H. (1984). The refined 2.2 Å (0.22 mm) X-ray crystal structure of the ternary complex formed by bovine trypsinogen, valine-valine, and the Arg 15 analogue of bovine pancreatic trypsin inhibitor. *Eur. J. Biochem. 144*, 185–190.
Bode, W., Wei, A.Z., Huber, R., Meyer, E., Travis, J., & Neumann, S. (1986). X-ray crystal structure of the complex of human leukocyte elastase (PMN elastase) and the third domain of the turkey ovomucoid inhibitor. *EMBO J. 5*, 2453–2458.
Boissel, J.P., LeBonniec, B., Rabiet, M.J., Labie, D., & Elion, J. (1984). Covalent structures of β and γ autolytic derivatives of human α-thrombin. *J. Biol. Chem. 259*, 5691–5697.
Brandstetter, H., Turk, D., Hoeffken, W., Grosse, D., Stürzebecher, J., Martin, P.D., Edwards, B.F.P., & Bode, W. (1992). X-ray crystal structure of thrombin complexes with the benzamidine- and arginine-based inhibitors NAPAP, 4-TAPAP and MQPA: A starting point for elaborating improved antithrombotics. *J. Mol. Biol.* (in press).
Braun, P.J., Hofsteenge, J., Chang, J.Y., & Stone, S.R. (1988). Preparation and characterization of proteolyzed forms of human α-thrombin. *Thromb. Res. 50*, 273–283.
Brezniak, D.V., Brower, M.S., Witting, J.I., Walz, D.A., & Fenton, J.W., II (1990). Human α- to ζ-thrombin cleavage occurs with neutrophil cathepsin G or chymotrypsin while fibrinogen clotting activity is retained. *Biochemistry 29*, 3526–3542.
Brower, M.S., Walz, D.A., Garry, E.E., & Fenton, J.W., II (1987). Thrombin-induced elastase alters human α-thrombin function: Limited proteolysis near the γ-cleavage site results in decreased fibrinogen clotting and platelet stimulatory activity. *Blood 69*, 813–819.
Brünger, A.T., Kuriyan, K., & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science 235*, 458–460.
Burley, S.K. & Petsko, G.A. (1985). Aromatic–aromatic interaction: A mechanism of protein structure stabilization. *Science 229*, 23–28.
Butkowski, R.J., Elion, J., Downing, M.R., & Mann, K.G. (1977). Primary structure of human prethrombin 2 and α-thrombin. *J. Biol. Chem. 252*, 4942–4957.
Cardin, A.D. & Weintraub, H.J.R. (1989). Molecular modeling of protein–glucosaminoglycan interactions. *Arteriosclerosis 9*, 21–32.
Carney, D.E., Herbosa, G.J., Stiernberg, J., Bergmann, J.S., Gordon, E.A., Scott, D.L., & Fenton, J.W., II (1986). Double signal hypothesis for thrombin initiation of cell proliferation. *Semin. Thromb. Hemostasis 12*, 231–240.
Chang, J.Y. (1986). The structures and proteolytic specificities of autolysed human thrombin. *Biochem. J. 240*, 797–802.
Chang, J.-Y. (1989). The hirudin-binding site of human α-thrombin; identification of lysyl residues which participate in the combining site of hirudin–thrombin complex. *J. Biol. Chem. 264*, 7141–7146.
Chang, T.L., Feinman, R.D., Laudis, B.H., & Fenton, J.W., II (1979). Antithrombin reactions with α- and γ-thrombins. *Biochemistry 18*, 113–119.
Charo, I.F., Bekeart, L.S., & Phillips, D.R. (1987). Platelet glycoprotein IIb-IIIa-like proteins mediate endothelial cell attachment to adhesive proteins and the extracellular matrix. *J. Biol. Chem. 262*, 9935–9938.
Chow, M.M., Meyer, E.F., Jr., Bode, W., Kam, C.-M., Radhakrishnan, R., Vijayalakshmi, J., & Powers, J.C. (1990). The 2.2 Å resolution X-ray crystal structure of the complex of trypsin inhibited by 4-chloro-3-ethoxy-7-guanidinocoumarin: A proposed model of the thrombin–inhibitor complex. *J. Am. Chem. Soc. 112*, 7783–7789.
Church, F.C., Pratt, C.W., Noyes, C.M., Kalayanamit, T., Sherril, G.B., Tobin, R.B., & Meade, B. (1989). Structural and functional properties of human α-thrombin, phosphopyridoxylated α-thrombin and γ_T-thrombin. Identification of lysyl residues in α-thrombin that are critical for heparin and fibrin(ogen) interactions. *J. Biol. Chem. 264*, 18419–18425.
Cohen, H.G., Silverton, E.W., & Davis, D. (1981). Refined crystal structure of γ-chymotrypsin at 1.9 Å resolution. Comparison with other pancreatic serine proteases. *J. Mol. Biol. 148*, 449–479.
Cruickshank, D.W.J. (1949). The accuracy of electron-density maps in X-ray analysis with special reference to dibenzyl. *Acta Crystallogr. 2*, 65–82.
Cunningham, D.D. & Farrell, D.H. (1986). Thrombin interactions with cultured fibroblasts: Relationship to mitogenic stimulation. *Ann. N.Y. Acad. Sci. 485*, 240–248.
De Cristofaro, R. & DiCera, E. (1990). Effect of protons on the amidase activity of human α-thrombin. *J. Mol. Biol. 216*, 1077–1085.
Degen, S.J.F., MacGillivray, R.T.A., & Davie, E.W. (1983). Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin. *Biochemistry 87*, 2087–2097.
Dihanich, M. & Monard, M. (1990). cDNA sequence of rat prothrombin. *Nucleic Acids Res. 18*, 4251.
Dodt, J., Köhler, S., & Baici, A. (1988). Interaction of site specific hirudin variants with α-thrombin. *FEBS Lett. 229*, 87–90.
Doyle, M.F. & Mann, K.G. (1990). Multiple active forms of thrombin IV. Relative activities of meizothrombins. *J. Biol. Chem. 265*, 10693–10710.
Elion, J., Boissel, J.P., LeBonniec, B., Bezeaud, A., Jandrot-Perrus, M., Rabiet, M.J., & Guillin, M.C. (1986). Proteolytic derivatives of thrombin. *Ann. N.Y. Acad. Sci. 485*, 16–26.
Esmon, N.L., Carrol, R.C., & Esmon, C.T. (1983). Thrombomodulin blocks the ability of thrombin to activate platelets. *J. Biol. Chem. 258*, 12238–12242.
Esmon, N.L., Owen, W.G., & Esmon, C.T. (1982). Isolation of membrane-bound cofactor for thrombin-catalyzed activation of protein C. *J. Biol. Chem. 257*, 859–864.
Fenton, J.W., II (1981). Thrombin specificity. *Ann. N.Y. Acad. Sci. 370*, 468–495.
Fenton, J.W., II (1986). Thrombin. *Ann. N.Y. Acad. Sci. 485*, 5–15.
Fenton, J.W., II (1988). Regulation of thrombin generation and functions. *Semin. Thromb. Hemostasis 14*, 234–240.
Fenton, J.W., II, Fasco, M.J., Stackrow, A.B., Aronson, D.L., Young, A.M., & Finlayson, J.S. (1977). Human thrombins: Production, evaluation and properties of α-thrombin. *J. Biol. Chem. 252*, 3587–3598.
Fenton, J.W., II, Olson, T.A., Zabinski, M.P., & Wilner, G.D. (1988). Anion binding exosite of human α-thrombin and fibrin(ogen) recognition. *Biochemistry 27*, 7106–7112.
Fenton, J.W., II, Witting, J.I., Pouliott, C., & Fareed, J. (1989). Thrombin anion-binding exosite interactions with heparin and various polyanions. *Ann. N.Y. Acad. Sci. 556*, 158–165.
Fenton, J.W., II, Zabinski, M.P., Hsieh, K., & Wilner, G.D. (1981). Thrombin non-covalent protein binding and fibrin(ogen) recognition. *Thromb. Haemostasis 46*, 177.
Folkers, P.J.M., Clore, G.M., Driscoll, P.C., Dodt, J., Kohler, S., & Gronenborn, A.M. (1989). Solution structure of recombinant hirudin and the Lys-47-Glu mutant: A nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing study. *Biochemistry 28*, 2601–2617.
Freer, S.T., Kraut, J., Robertus, J.D., Wright, H.T., & Xuong, N.H. (1970). Chymotrypsinogen: 2.5 Å Crystal structure, comparison with α-chymotrypsin, and implications for zymogen activation. *Biochemistry 9*, 1997–2009.
Friezner-Degen, S.J., Schaefer, L.A., Jamison, C.S., Grant, S.G., Fitzgibbon, J.J., Pai, J.A., Chapman, V.M., & Elliott, R.W. (1990). Characterization of the cDNA coding for mouse prothrombin and localization of the gene on mouse chromosome 2. *DNA Cell Biol. 9*, 487–498.
Gilson, M.K., Rashin, A., Fine, R., & Honig, B. (1985). On the calculation of electrostatic interactions in proteins. *J. Mol. Biol. 183*, 503–516.

Glenn, K.C., Frost, G.H., Bergmann, J.S., & Carney, D.H. (1988). Synthetic peptides bind to high-affinity thrombin receptors and modulate thrombin mitogenesis. *Peptide Res. 1*, 65–73.

Greer, J. (1981). Comparative model-building of the mammalian serine proteinases. *J. Mol. Biol. 153*, 1027–1042.

Griffith, M.J. (1982). Kinetics of the heparin-enhanced antithrombin III/thrombin reaction. Evidence for a template model for the mechanism of action of heparin. *J. Biol. Chem. 257*, 7360–7365.

Griffith, M.J., Kingdon, H.S., & Lundblad, R.L. (1979). The interaction of heparin with human α-thrombin: Effect on the hydrolysis of anilide tripeptide substrates. *Arch. Biochem. Biophys. 195*, 378–384.

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 inhibitor. *EMBO J. 9*, 2361–2365.

Hageman, T.C., Endres, G.F., & Scheraga, H.A. (1975). Mechanism of action of thrombin on fibrinogen; on the role of the A chain of bovine thrombin in specificity and differentiating between thrombin and trypsin. *Arch. Biochem. Biophys. 171*, 327–336.

Hageman, T.C. & Scheraga, H.A. (1974). Mechanism of action of thrombin on fibrinogen; reaction of the *N*-terminal CNBR fragment from the Aα chain of human fibrinogen with bovine thrombin. *Arch. Biochem. Biophys. 164*, 707–715.

Hartley, B.S. & Kauffman, D. (1966). Corrections to the amino acid sequence of bovine chymotrypsinogen A. *Biochem. J. 101*, 229–231.

Hartley, B.S. & Shotton, D.M. (1971). Pancreatic elastase. In *Enzymes*, Vol. 3 (Boyer, P.D., Ed.), pp. 323–373. Academic Press, New York, London.

Haruyama, H. & Wüthrich, K. (1989). The conformation of recombinant disulfatohirudin in aqueous solution determined by nuclear magnetic resonance. *Biochemistry 28*, 4301–4312.

Henriksen, R.A. & Mann, K.G. (1988). Identification of the primary structural defect in the dysthrombin Thrombin Quick I: Substitution of cysteine for arginine-382. *Biochemistry 27*, 9160–9165.

Henriksen, R.A. & Mann, K.G. (1989). Substitution of valine for glycine-558 in the congenital dysthrombin Thrombin Quick II alters primary substrate specificity. *Biochemistry 28*, 2078–2082.

Heuck, C.C., Schiele, V., Horn, D., Fronda, D., & Ritz, E. (1985). The role of surface charge on the accelerating action of heparin on the antithrombin III-inhibited activity of α-thrombin. *J. Biol. Chem. 260*, 4598–4603.

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

Hofsteenge, J. & Stone, S.R. (1987). The effect of thrombomodulin on the cleavage of fibrinogen fragments by thrombin. *Eur. J. Biochem. 168*, 49–56.

Hofsteenge, J., Taguchi, H., & Stone, S.R. (1986). Effect of thrombomodulin on the kinetics of the interaction of thrombin with substrates and inhibitors. *Biochem. J. 237*, 243–251.

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

Hogg, P.J. & Jackson, C.M. (1989). Fibrin monomer protects thrombin from inactivation by heparin-antithrombin III: Implications for heparin efficacy. *Proc. Natl. Acad. Sci. USA 86*, 3619–3623.

Horne, McD.K. & Gralnick, H.R. (1983). The oligosaccharide of human thrombin: Investigations of functional significance. *Blood 63*, 188–194.

Huber, R. & Bode, W. (1978). Structural basis of the activation and action of trypsin. *Acc. Chem. Res. 11*, 114–122.

Huber, R. & Carrell, R.W. (1989). Implications of the three-dimensional structure of α_1-antitrypsin for structure and function of serpins. *Biochemistry 28*, 8951–8966.

Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., & Steigemann, W. (1974). Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. II. Crystallographic refinement at 1.9 Å resolution. *J. Mol. Biol. 89*, 73–101.

Huber, R. & Schneider, M. (1985). A group refinement procedure in protein crystallography using Fourier transforms. *J. Appl. Crystallogr. 18*, 165–169.

Jack, A. & Levitt, M. (1978). Refinement of large structures by simultaneous minimization of energy and R-factor. *Acta Crystallogr. A 34*, 931–935.

Jackson, C.M. & Nemerson, Y. (1980). Blood coagulation. *Annu. Rev. Biochem. 49*, 765–811.

Jakubowski, H.V. & Owen, W.G. (1989). Macromolecular specificity determinants on thrombin for fibrinogen and thrombomodulin. *J. Biol. Chem. 264*, 11117–11121.

Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr. 15*, 23–31.

Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: Pattern, recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*, 2577–2637.

Kaczmarek, E. & McDonagh, J. (1988). Thrombin binding to the Aα-, Bβ- and γ-chains of fibrinogen and to their remnants contained in fragment E. *J. Biol. Chem. 263*, 13896–13900.

Kam, C.M., Fujikawa, W., & Powers, J.C. (1988). Mechanism-based isocoumarin inhibitors for trypsin and blood coagulation serine proteases: New anticoagulants. *Biochemistry 27*, 2547–2557.

Kaminski, M. & McDonagh, J. (1987). Inhibited thrombins; interactions with fibrinogen and fibrin. *Biochem J. 242*, 881–887.

Karpatkin, S. & Karpatkin, M. (1974). Inhibition of the enzymatic activity of thrombin by concanavalin A. *Biochem. Biophys. Res. Commun. 57*, 1111–1118.

Karshikov, A.D., Engh, R., Bode, W., & Atanasov, B.P. (1989). Electrostatic interactions in proteins: Calculations of the electrostatic term of free energy and the electrostatic potential field. *Eur. Biophys. J. 17*, 287–297.

Kawabata, S., Morita, T., Iwanaga, S., & Igarashi, H. (1985). Staphylocoagulase-binding region in human prothrombin. *J. Biochem. 97*, 325–331.

Kettner, C., Mersinger, L., & Knabb, R. (1991). The selective inhibition of thrombin by peptides of boroarginine. *J. Biol. Chem. 265*, 18289–18297.

Kettner, C. & Shaw, E. (1979). D-Phe-Pro-Arg CH_2Cl—A selective affinity label for thrombin. *Thromb. Res. 14*, 969–973.

Kettner, C. & Shaw, E. (1981). Inactivation of trypsin-like enzymes with peptides of arginine chloromethyl ketone. *Methods Enzymol. 80*, 826–842.

Klapper, I., Hagstrom, R., Fine, R., Sharp, K., & Honig, B. (1986). Focusing of electric fields in the active site of Cu–Zn superoxide dismutase: Effect of ionic strength and amino-acid modification. *Proteins Struct. Funct. Genet. 1*, 47–59.

Landis, C., Köhler, R.A., & Fenton, J.W., II (1981). Human thrombins. *J. Biol. Chem. 256*, 4604–4610.

Laskowski, M. & Kato, I. (1980). Protein inhibitors of proteinases. *Annu. Rev. Biochem. 49*, 593–626.

LeBonniec, B.F. & Esmon, C.T. (1991). Glu-192 → Gln substitution in thrombin mimics the catalytic switch induced by thrombomodulin. *Proc. Natl. Acad. Sci. USA 88*, 7371–7375.

Le Bonniec, B.F., MacGillivray, R.T.A., & Esmon, C.T. (1991). Thrombin Glu-39 restricts the P'3 specificity to nonacidic residues. *J. Biol. Chem. 266*, 13796–13803.

Lee, B. & Richards, F.M. (1971). The interpretation of protein structure: Estimation of static accessibility. *J. Mol. Biol. 55*, 379–400.

Levitt, M. (1974). Energy refinement of egg-white lysozyme. *J. Mol. Biol. 82*, 393–420.

Lewis, S.D., Lorand, L., Fenton, J.W., II, & Shafer, J.A. (1987). Catalytic competence of human α- and γ-thrombin in the activation of fibrinogen and factor XIII. *Biochemistry 26*, 7597–7603.

Li, E.H.H., Fenton, J.W., II, & Feinman, R.D. (1976). The role of heparin in the thrombin–antithrombin III reaction. *Arch. Biochem. Biophys. 175*, 153–159.

Liem, R.K.H. & Scheraga, H.A. (1974). Mechanism of action of thrombin on fibrinogen N. Further mapping of the active sites of thrombin and trypsin. The binding of thrombin and fibrin. *Arch. Biochem. Biophys. 160*, 333–339.

Liu, C.Y., Nossel, H.L., & Kaplan, K.L. (1979). The binding of thrombin by fibrin. *J. Biol. Chem. 254*, 10421–10425.

Liu, L.W., Vu, T.K.H., Esmon, C.T., & Coughlin, S.R. (1991). The region of the thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. *J. Biol. Chem. 266*, 16977–16980.

Löbermann, H., Tokuoka, R., Deisenhofer, J., & Huber, R. (1984). Human α_1-proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. *J. Mol. Biol. 177*, 531–556.

Lottenberg, R., Hall, J.A., Blinder, M., Binder, E.P., & Jackson, C.M. (1983). The action of thrombin on peptide *p*-nitroanilide substrates.

Substrate selectivity and examination of hydrolysis under different reaction conditions. *Biochim. Biophys. Acta 742*, 539-557.

Lundblad, R.L., Noyes, C.M., Featherstone, G.L., Harrison, J.H., & Jenzano, J.W. (1988). The reaction of bovine α-thrombin with tetranitromethane. Characterization of the modified protein. *J. Biol. Chem. 263*, 3729-3734.

Lundblad, R.L., Noyes, C.M., Mann, K.G., & Kingdon, H.S. (1979). The covalent differences between bovine α- and β-thrombin; a structural explanation for the changes in catalytic activity. *J. Biol. Chem. 254*, 8524-8528.

Luzzati, V. (1952). Traitement statistique des erreurs dans la determination des structures crystallines. *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., Peterson, T.E., Sottrup-Jensen, L., & Claeys, H. (1975). Complete primary structure of prothrombin: Isolation, structure and reactivity of ten carboxylated glutamic acid residues and regulation of prothrombin activation by thrombin. In *Proteases and Biological Control* (Reich, E., Rifkin, D.B., & Shaw, E., Eds.), pp. 123-149. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Mao, S.J.T., Yates, M.T., Owen, T.J., & Krstenansky, J.L. (1988). Interaction of hirudin with thrombin: Identification of a minimal binding domain of hirudin that inhibits clotting activity. *Biochemistry 27*, 8170-8173.

Maraganore, J.M., Bourdon, P., Jablonski, J., Ramachandran, K.L., & Fenton, J.W., II (1990). Design and characterization of hirulogs: A novel class of bivalent peptide inhibitors of thrombin. *Biochemistry 29*, 7095-7101.

Marsh, H.C., Meinwald, Y.C., Lee, S., Martinelli, R.A., & Scheraga, H.A. (1985). Mechanism of action of thrombin on fibrinogen. Direct evidence for the involvement of phenylalanine at position P9. *Biochemistry 24*, 2806-2812.

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

Matsuzaki, T., Sasaki, C., Okumura, C., & Umeyama, H. (1989). X-ray analysis of a thrombin inhibitor-trypsin complex. *J. Biochem. (Tokyo) 105*, 949-952.

Matthew, J.B. (1985). Electrostatic effects in proteins. *Annu. Rev. Biophys. Biochem. 14*, 387-417.

McGowan, E.B. & Detwiler, T.C. (1986). Modified platelet responses to thrombin. Evidences for two types of receptors or coupling mechanisms. *J. Biol. Chem. 261*, 739-746.

Meinwald, Y.C., Martinelli, R.A., Van Nispen, J.W., & Scheraga, H.A. (1980). Mechanism of action of thrombin on fibrinogen. Size of the Aα fibrinogen-like peptide that contacts the active site of thrombin. *Biochemistry 19*, 3820-3825.

Meloun, B., Kluh, I., Kostka, V., Moravek, L., Prusik, Z., Vanecek, J., Keil, B., & Sorm, F. (1966). Covalent structure of bovine chymotrypsin A. *Biochim. Biophys. Acta, 130*, 543-546.

Messerschmidt, A. & Pflugrath, J.W. (1987). Crystal orientation and X-ray pattern prediction routines for area detector diffractometer systems in macromolecular crystallography. *J. Appl. Crystallogr. 20*, 306-315.

Messerschmidt, A., Schneider, M., & Huber, R. (1990). ABSCOR: A scaling and absorption correction program for the FAST area detector diffractometer. *J. Appl. Crystallogr. 23*, 436-439.

Meyer, E., Cole, G., Radhakrishnan, R., & Epp, O. (1988). Structure of native porcine pancreatic elastase at 1.65 Å resolution. *Acta Crystallogr. B44*, 26-38.

Mikes, O., Holeysovsky, V., Tomasek, V., & Sorm, F. (1966). Covalent structure of bovine trypsinogen. The position of the remaining amides. *Biochem. Biophys. Res. Commun. 24*, 346-352.

Miyata, T., Morita, T., Inomoto, T., Kawauchi, S., Shirakami, A., & Iwanaga, S. (1987). Prothrombin Tokushima, a replacement of arginine-418 by tryptophan that impairs the fibrinogen clotting activity of derived thrombin Tokushima. *Biochemistry 26*, 1117-1122.

Naski, M.C., Fenton, J.W., II, Maraganore, J.M., Olson, S.T., & Shafer, J.A. (1990). The COOH-terminal domain of hirudin. An exosite-directed competitive inhibitor of the action of α-thrombin on fibrinogen. *J. Biol. Chem. 265*, 13484-13489.

Needleman, S.B. & Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequences of two proteins. *J. Mol. Biol. 48*, 443-453.

Nesheim, M.E. (1983). A simple rate law that describes the kinetics of the heparin-catalyzed reaction between antithrombin III and thrombin. *J. Biol. Chem. 258*, 14708-14717.

Ni, F., Konishi, Y., Frazier, R.B., & Scheraga, H.A. (1989a). High-resolution NMR studies of fibrinogen-like peptides in solution: Interaction of thrombin with residues 1-23 of the Aα chain of human fibrinogen. *Biochemistry 28*, 3082-3094.

Ni, F., Meinwald, Y.C., Vasquez, M., & Scheraga, H.A. (1989b). Structure of a thrombin-bound peptide corresponding to residues 7-16 of the Aα-chain of human fibrinogen. *Biochemistry 28*, 3094-3105.

Nilsson, B., Horne, M.K., & Gralnick, H.R. (1983). The carbohydrate of human thrombin: Structural analysis of glycoprotein oligosaccharides by mass spectrometry. *Arch. Biochem. Biophys. 224*, 127-133.

Noé, G., Hofsteenge, J., Rovelli, G., & Stone, S.R. (1988). The use of sequence specific antibodies to identify a secondary binding site in thrombin. *J. Biol. Chem. 263*, 11729-11735.

Okamoto, S., Hijikata, A., Kikumoto, R., Tonomura, S., Hara, N., Ninomiya, K., Maruyama, A., Sugano, M., & Tamao, Y. (1981). Potent inhibition of thrombin by the newly synthesized arginine derivative no. 805. The importance of stereostructure of its hydrophobic carboxamide portion. *Biochem. Biophys. Res. Commun. 101*, 440-446.

Olson, S.T. & Shore, J.D. (1982). Demonstration of a two-step reaction mechanism for inhibition of α-thrombin by antithrombin III and identification of the step affected by heparin. *J. Biol. Chem. 257*, 14891-14895.

Olson, T.A., Sonder, S.A., Wilner, G.D., & Fenton, J.W., II (1986). Heparin binding in proximity to the catalytic site of human α-thrombin. *Ann. N.Y. Acad. Sci. 485*, 96-103.

Patthy, L. (1985). Evolution of proteases of blood coagulation and fibrinolysis by assembly of modules. *Cell 41*, 657-663.

Pflugrath, J.W., Saper, M.A., & Quiocho, F.A. (1984). In *Methods and Application in Crystallographic Computing* (Hall, S. & Ashiaka, T., Eds.), pp. 404-407. Clarendon Press, Oxford.

Pomerantz, M.W. & Owen, W.G. (1978). A catalytic role for heparin. Evidence for a ternary complex of heparin cofactor thrombin and heparin. *Biochim. Biophys. Acta 535*, 66-77.

Preissner, K.T., DelVos, V., & Müller-Berghaus, G. (1987). Binding of thrombin to thrombomodulin accelerates inhibition of the enzyme by antithrombin III. Evidence for a heparin-independent mechanism. *Biochemistry 26*, 2521-2528.

Prescott, S.M., Seeger, A.R., Zimmerman, G.A., McIntyre, T.M., & Maraganore, J.M. (1990). Hirudin-based peptides block the inflammatory effects of thrombin on endothelial cells. *J. Biol. Chem. 265*, 9614-9616.

Priestle, J.P. (1988). RIBBON: A stereo cartoon drawing program for proteins. *J. Appl. Crystallogr. 21*, 572-576.

Ramakrishnan, C. & Ramachandran, G.N. (1965). Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys. J. 5*, 909-933.

Read, R.J. & James, M.N.G. (1986). Introduction to the proteinase inhibitors: X-ray crystallography. In *Proteinase Inhibitors* (Barret, A.J. & Salvesen, G., Eds.), pp. 301-336. Elsevier, Amsterdam.

Remington, S.J., Woodbury, R.G., Reynolds, R.A., Matthews, B.W., & Neurath, H. (1988). The structure of rat mast cell protease II at 1.9 Å resolution. *Biochemistry 27*, 8097-8105.

Richards, F.M. (1977). Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng. 6*, 151-170.

Richardson, J.S. & Richardson, D.C. (1990). Principles and patterns of protein conformation. In *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G., Ed.), pp. 1-98. Plenum Press, New York.

Rosenberg, R.D. & Damus, P.S. (1973). The purification and mechanism of action of human antithrombin-heparin cofactor. *J. Biol. Chem. 248*, 6490-6505.

Rossmann, M.G. & Argos, P. (1975). A comparison of the heme binding pocket in globins and cytochrome b5. *J. Biol. Chem. 250*, 7525-7532.

Rydel, T.J., Ravichandran, K.G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., & Fenton, J.W., II (1990). The structure of a complex of recombinant hirudin and human α-thrombin. *Science 249*, 277-280.

Rydel, T.J., Tulinsky, A., Bode, W., & Huber, R. (1991). Refined structure of the hirudin-thrombin complex. *J. Mol. Biol. 221*, 583–601.

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

Scheraga, H.A. (1977). Active site mapping of thrombin. In *Chemistry and Biology of Thrombin* (Lundblad, R.L., Fenton, J.W., II, & Mann, K.G., Eds.), pp. 145–158. Ann Arbor Science Publ., Ann Arbor, Michigan.

Schirmer, T., Bode, W., & Huber, R. (1987). Refined three-dimensional structure of two cyanobacterial C-phycocyanins at 2.1 and 2.5 Å resolution. *J. Mol. Biol. 196*, 677–695.

Schuman, M.A. (1986). Thrombin – Cellular interactions. *Ann. N.Y. Acad. Sci. 485*, 228–239.

Skaug, K. & Christenson, T.B. (1971). The significance of the carbohydrate constituents of bovine thrombin for the clotting activity. *Biochim. Biophys. Acta 230*, 627–629.

Sonder, S.A. & Fenton, J.W., II (1984). Proflavin binding within the fibrinopeptide groove adjacent to the catalytic site of human α-thrombin. *Biochemistry 23*, 1818–1823.

Steigemann, W. (1974). Die Entwicklung und Anwendung von Rechenverfahren und Rechenprogrammen zur Strukturanalyse von Proteinen am Beispiel des Trypsin-Trypsininhibitor Komplexes, des freien Inhibitors und der L-Asparaginase. Ph.D. Thesis, University of Munich.

Stone, S.R., Braun, P.J., & Hofsteenge, J. (1987). Identification of regions of α-thrombin involved in its interaction with hirudin. *Biochemistry 26*, 4617–4624.

Stone, S.R. & Hofsteenge, J. (1986). Kinetics of the inhibition of thrombin by hirudin. *Biochemistry 25*, 4622–4628.

Stubbs, M., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S.R., & Bode, W. (1992). The interaction of thrombin with fibrinogen – A structural basis for its specificity. *Eur. J. Biochem.* (in press).

Stürzebecher, J., Markwardt, F., Voigt, B., Wagner, G., & Walsmann, P. (1983). Cyclic amides of Nα-arylsulfonylaminoacylated 4-amidinophenylalanine – Tight binding of thrombin. *Thromb. Res. 29*, 635–642.

Sugawara, Y., Birktoft, J.J., & Berliner, L.J. (1986). Human α- and γ-thrombin inhibition by trypsin inhibitors supports predictions from molecular graphics experiments. *Semin. Thromb. Hemostasis 12*, 209–210.

Suzuki, K., Nishioka, J., & Hayashi, T. (1990). Localization of thrombomodulin-binding site within human thrombin. *J. Biol. Chem. 265*, 13263–13267.

Svendsen, L., Blombäck, B., Blombäck, M., & Olsson, P.I. (1972). Synthetic chromogenic substrates for determination of trypsin, thrombin and thrombin-like enzymes. *Thromb. Res. 1*, 267–278.

Tanford, D. & Kirkwood, J.G. (1957). Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc. 79*, 5333–5339.

Thomas, K.A., Smith, G.M., Thomas, T.B., & Feldmann, R.J. (1982). Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. *Proc. Natl. Acad. Sci. USA 79*, 4843–4847.

Thompson, A.R. (1976). High affinity binding of human and bovine thrombin to *p*-chlorobenzylamido-ε-aminocaproyl-agarose. *Biochim. Biophys. Acta 422*, 200–209.

Tollefsen, D.M., Majerus, D.W., & Blank, M.K. (1982). Heparin cofactor II. Purification and properties of a heparin-dependent inhibitor of thrombin in human plasma. *J. Biol. Chem. 257*, 2162–2169.

Toma, K. & Suzuki, K. (1989). Mapping active sites of blood coagulation serine proteinases – activated protein C and thrombin – on simple graphics models. *J. Mol. Graphics 7*, 146–149.

Tsiang, M., Lentz, S.R., Dittman, W.A., Scarpati, E.M., & Sadler, J.E. (1990). Equilibrium binding of thrombin to recombinant human thrombomodulin: Effect of hirudin, fibrinogen, factor Va and peptide analogues. *Biochemistry 29*, 10602–10612.

Tsukada, H. & Blow, D. (1985). Structure of α-chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol. 184*, 703–711.

Turk, D., Stürzebecher, J., & Bode, W. (1991). Geometry of binding of α-tosylated piperidides of *m*-amidino, *p*-amidino- and *p*-guanidino phenylalanine to thrombin and trypsin. X-ray crystal structures of their trypsin complexes and modeling of their thrombin complexes. *FEBS Lett. 287*, 133–138.

Vali, Z. & Scheraga, H.A. (1988). Localization of the binding site on fibrin for the secondary binding site of thrombin. *Biochemistry 27*, 1956–1963.

Van Nispen, J.W., Hageman, T.C., & Scheraga, H.A. (1977). Mechanism of action of thrombin on fibrinogen; the reaction of thrombin with fibrinogen-like peptides containing 11, 14 and 16 residues. *Arch. Biochem. Biophys. 182*, 227–243.

Vu, T.K.H., 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.

Wallace, A., Rovelli, G., Hofsteenge, J., & Stone, S.R. (1989). Effect of heparin on the glia-derived-nexin–thrombin interaction. *Biochem. J. 257*, 191–196.

Walsh, K.A. & Neurath, H. (1964). Trypsinogen and chymotrypsinogen as homologous proteins. *Proc. Natl. Acad. Sci. USA 52*, 884–889.

Walsmann, P. & Markwardt, F. (1981). Biochemische und pharmakologische Aspekte des Thrombininhibitors Hirudin. *Pharmazie 36*, 653–660.

Wang, D., Bode, W., & Huber, R. (1985). Bovine chymotrypsinogen A. X-ray crystal structure analysis & refinement of a new crystal form at 1.8 resolution. *J. Mol. Biol. 185*, 595–624.

Wei, A.-Z., Mayr, I., & Bode, W. (1988). The refined 2.3 Å crystal structure of human leukocyte elastase in a complex with a valine chloromethylketone inhibitor. *FEBS Lett. 234*, 367–373.

White, G.C., Lundblad, R.L., & Griffith, M.J. (1981). Structure–function relations in platelet–thrombin reactions. *J. Biol. Chem. 256*, 1763–1766.

Wu, Q., Sheehan, J.P., Tsiang, M., Lentz, S.R., Birktoft, J.J., & Sadler, J.E. (1991). Single amino acid substitutions dissociate fibrinogen-clotting and thrombomodulin-binding activities of human thrombin.