The EMBO Journal vol.14 no.21 pp.5149–5157, 1995

Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin

Andreas van de Locht[1], Doriano Lamba, Margit Bauer, Robert Huber, Thomas Friedrich[2], Burkhard Kröger[2], Wolfgang Höffken[2] and Wolfram Bode[1]

Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, 82152 Martinsried and [2]Abteilung für Biotechnologie, BASF Aktiengesellschaft, 67056 Ludwigshafen, Germany

[1]Corresponding authors

Rhodniin is a highly specific inhibitor of thrombin isolated from the assassin bug *Rhodnius prolixus*. The 2.6 Å crystal structure of the non-covalent complex between recombinant rhodniin and bovine α-thrombin reveals that the two Kazal-type domains of rhodniin bind to different sites of thrombin. The amino-terminal domain binds in a substrate-like manner to the narrow active-site cleft of thrombin; the imidazole group of the P1 His residue extends into the S1 pocket to form favourable hydrogen/ionic bonds with Asp189 at its bottom, and additionally with Glu192 at its entrance. The carboxy-terminal domain, whose distorted reactive-site loop cannot adopt the canonical conformation, docks to the fibrinogen recognition exosite via extensive electrostatic interactions. The rather acidic polypeptide linking the two domains is displaced from the thrombin surface, with none of its residues involved in direct salt bridges with thrombin. The tight (K_i = 2×10^{-13} M) binding of rhodniin to thrombin is the result of the sum of steric and charge complementarity of the amino-terminal domain towards the active-site cleft, and of the electrostatic interactions between the carboxy-terminal domain and the exosite.

Keywords: crystal structure/electrostatic interaction/hirudin/Kazal-type/thrombin inhibitor

Introduction

α-Thrombin (EC 3.4.21.5) is a serine proteinase of trypsin-like specificity, which is activated and liberated into the bloodstream *in vivo* by limited factor Xa cleavage of its inactive precursor prothrombin. α-Thrombin, the key enzyme in hemostasis and thrombosis, exhibits both enzymatic and hormone-like properties, and can be both pro- and anti-coagulatory. It cleaves fibrinogen to fibrin and stimulates platelet aggregation via cleavage of the thrombin receptor, in this way effecting thrombus formation. Furthermore, it activates other coagulation factors and interacts with a variety of other cells. Thrombin activity *in vivo* is regulated by the plasma inhibitors α_2-macroglobulin and by the serpins antithrombin III, heparin cofactor II and proteinase nexin I (for reviews see Stubbs and Bode, 1993; Kalafatis *et al.*, 1994).

Several characteristic structural features enable thrombin to accomplish these divergent functions with high specificity. Its canyon-like active-site cleft is narrowed by two large insertion loops, the rigid 60 loop and the more flexible 149 loop opposing it [thrombin residues are identified via the chymotrypsinogen numbering deduced from topological equivalence (Bode *et al.*, 1989, 1992); the suffix 'I' serves to distinguish rhodniin from thrombin residues]. This appears to restrict access of macromolecular substrates and inhibitors to the catalytic residues at its base (Bode *et al.*, 1989, 1992). In contrast to the digestive enzyme trypsin, thrombin prefers Arg to Lys residues at the P1 position of substrates (P1, P2, etc. and P1′, P2′, etc. designate substrate residue positions amino- and carboxy-terminal of the scissile peptide bond respectively, and S1, S2 and S1′, S2′, etc. the corresponding subsites in the proteinase) (Lottenberg *et al.*, 1983). Another characteristic feature of α-thrombin is the clustering of like charges giving rise to a sandwich-like electric field distribution (Bode *et al.*, 1992). The active-site residues are at the edge of a negatively charged surface area. A positively charged surface patch situated at the carboxy-terminal exit of the active-site cleft (called the 'fibrinogen recognition exosite') interacts with negatively charged groups of many specific thrombin substrates and inhibitors (Hofsteenge *et al.*, 1986), and another positive patch has been implicated in heparin binding (Church *et al.*, 1989; Ye *et al.*, 1994).

In order to feed on blood from mammalian hosts, hematophageous animals must prevent local clotting of the victim's blood. These parasites have developed various anti-clotting mechanisms (for a review see Markwardt, 1994), including the specific inhibition of thrombin. The most prominent of these natural thrombin-directed inhibitors is hirudin, a 65 residue polypeptide isolated from the European leech *Hirudo medicinalis* (Walsmann and Markwardt, 1981; Stone and Hofsteenge, 1986; Dodt *et al.*, 1988). The specificity and potency of this inhibitor results from the combined interaction of the compact main body with thrombin surface sites close to the active site, and of the acidic carboxy-terminal tail with the basic fibrinogen recognition exosite (Rydel *et al.*, 1990, 1991; Grütter *et al.*, 1990). Specific thrombin inhibitors with hirudin-like (Scacheri *et al.*, 1993) or other sequences (Steiner *et al.*, 1992; Strube *et al.*, 1993) have been isolated from other leeches and also from snake venom (Zingali *et al.*, 1993).

Anti-coagulatory agents from blood sucking insects have also been identified (see Markwardt, 1994) including platelet aggregation inhibitors, factor Xa inhibitors and the 103 residue thrombin inhibitor rhodniin from *Rhodnius prolixus* (Friedrich *et al.*, 1993). Rhodniin is composed of two Kazal-type domains and exhibits an inhibition constant of ~2×10^{-13} M. Kazal-type inhibitors had previously been

Table I. Alignment of rhodniin's amino and carboxy-terminal domain with the third domain of the turkey ovomucoid inhibitor (OMTKY3) (Kato *et al.*, 1978)

	1 5 10 15 20 25 29
OMTKY3	· L A A V S V D C S E Y P K P A C T L E Y R P L C G S D N K
	1I 5I 10I 15I 20I
Rhod1	E G G E P C A C P H A L H R V C G S D G E
	49I 55I 60I 65I 70I
Rhod2	E P D E D E D V C Q E C D G D E Y K P V C G S D D I
	30 35 40 45 50 56
OMTKY3	T Y G N K C N F C N A V V E S N G T L T L S H F G K C
	25I 30I 35I 40I 45I 48I
Rhod1	T Y S N P C T L N C A K F N G K P E L V K V H D G P C
	75I 80I 85I 90I 95I 100I 103I
Rhod2	T Y D N N C R L E C A S I S S S P G V E L K H E G P C R T

The sequences were aligned due to structural homology. 'Rhod1' and 'Rhod2' denote rhodniin's amino- and carboxy-terminal domain respectively. Compared with classical Kazal-type inhibitors the amino-terminal part of both rhodniin domains is truncated to one and two residues respectively, between the first and the second cysteine residue. In the carboxy-terminal domain one amino acid is inserted between the second and the third cysteine. Cysteine residues are marked by boxes, His10I at position P1 is displayed inverse. The aligned sequences were formatted using the program ALSCRIPT (Barton, 1993).

suggested as candidates for thrombin inhibition based on modelling studies (Bode *et al.*, 1992), since they would minimize steric hindrance with thrombin's active-site cleft. In spite of the early finding of the bovine pancreatic secretory trypsin inhibitor as an anti-coagulant protein (Kazal *et al.*, 1948), no Kazal-type inhibitors displayed strong inhibitory activity towards thrombin. Here we describe the crystal structure of the non-covalent complex between rhodniin and thrombin to elucidate the rhodniin structure and its mechanism of interaction with thrombin.

Results

The complex between rhodniin and thrombin

Rhodniin consists of two compact domains of essentially Kazal type, which are covalently connected through an acidic linker peptide (see Table I). The spatial structure shows residues 1I–48I to constitute the amino-terminal domain and residues 55I–103I, the carboxy-terminal domain. Each of the two domains binds independently towards completely separate surface sites of thrombin (see Figure 1). The amino-terminal domain of rhodniin interacts in a substrate-like manner with thrombin's active site via its reactive-site loop, whereas the carboxy-terminal domain, rotated by almost 90° along the connecting linker segment, faces the fibrinogen recognition exosite of thrombin with its flat side. The inter-domain linker shows an overall extended conformation and has only a few van der Waals contacts with the thrombin surface. The interface between the rhodniin molecule and thrombin which is buried upon complex formation comprises ~1900 Å^2, and as such is of similar size to that of the hirudin–thrombin complex. The first domain accounts for 63% of this interface, whereas the contribution of the second domain is 33%.

The amino-terminal domain

The amino-terminal domain of rhodniin resembles a thin disc. Its core comprises the typical secondary structure elements of the classical Kazal inhibitor (Weber *et al.*, 1981; Bolognesi *et al.*, 1982): i.e. a short β-hairpin loop (with strands 14I–17I and 20I–23I), followed by a three-turn α-helix (25I–35I) that, after an open turn, ends in a curved C-terminal segment, which aligns with the β-loop and forms a short three-stranded antiparallel β-pleated sheet (see Figure 2). The carboxy terminus is disulfide bridged to the first strand via Cys16I–Cys48I. The segment Ala7I(P4) to Arg14I(P4′) forms the characteristically exposed reactive-site loop, which on both sides of the scissile peptide bond, His10I–Ala11I, is anchored to the molecular core; through disulfide bridge Cys8I(P3)–Cys27I on the amino-terminal side, and on the carboxy-terminal side via hydrogen bonds formed between the amide nitrogen of the highly conserved Asn25I (Laskowski *et al.*, 1990) and the carbonyl groups of Pro9I(P2) and Ala11I(P1′).

In strong contrast to the classical Kazal-type inhibitors, however, which have a seven to eight amino acid spacer between the first and the second cysteine, in the first rhodniin domain these cysteine residues are separated by a single residue only. Nevertheless, the amino-terminal segment is still disulfide-clamped to the central helix without distortions at the reactive site or within the domain scaffold through a shift of the disulfide bridge partner of the first cysteine downstream by one position in the helix. This results in an almost perpendicular orientation of this Cys6I–Cys31I disulfide bridge compared with that in classical Kazal inhibitors. As a consequence, the amino-terminal segment Glu1I–Pro5I cannot run across the flat side, but is aligned along the small corner of the domain, so that the rhodniin domain exhibits a disc-like rather than a wedge-like shape, and the hydrophobic core becomes partially accessible.

Interaction of the amino-terminal domain with thrombin

The reactive-site loop of the first rhodniin domain from Cys6I(P5) to Arg14I(P4′) is in direct contact with thrombin and binds to its active site in a substrate-like manner, that is with the main-chain conformation and intermolecular hydrogen bond interactions characteristic of canonical binding proteinase inhibitors (Bode and Huber, 1992). The four residues preceding the scissile peptide bond are aligned to the enzyme segment Ser214–Gly216 in an antiparallel, twisted manner, forming the typical single (between Ser214 O and His10I N) and double hydrogen bonds (between Gly216 N,O and Cys8I O,N) (see Figure 3); due to a sizable kink at Ala7I(P4), an additional hydrogen bond is formed between Cys6I(P5) O and Gly219 N. In agreement with the canonical binding, the His10I(P1) carbonyl group of the scissile peptide bond is situated in the oxyanion hole fixed via a bifurcated hydrogen bond to Gly193 N and Ser195 N, and the following three residue segment runs antiparallel to thrombin segment Leu40–Leu41, forming the usual hydrogen bond between Leu12I(P2′) N and Leu41 O.

The imidazole side chain of the P1 residue His10I extends into the S1 pocket, leaving a gap to the Asp189 carboxylate group at the bottom of this pocket which is spanned by a buried solvent molecule. This water molecule is a hydrogen bond donor to both carboxylate oxygens (2.8 Å and 3.1 Å) and to Gly219 O (3.3 Å), and an acceptor from His10I N$^{\varepsilon 2}$. His10I(P1)N$^{\delta 1}$ donates another hydrogen bond to one carboxylate oxygen of Glu192 (2.8 Å). The second carboxy-

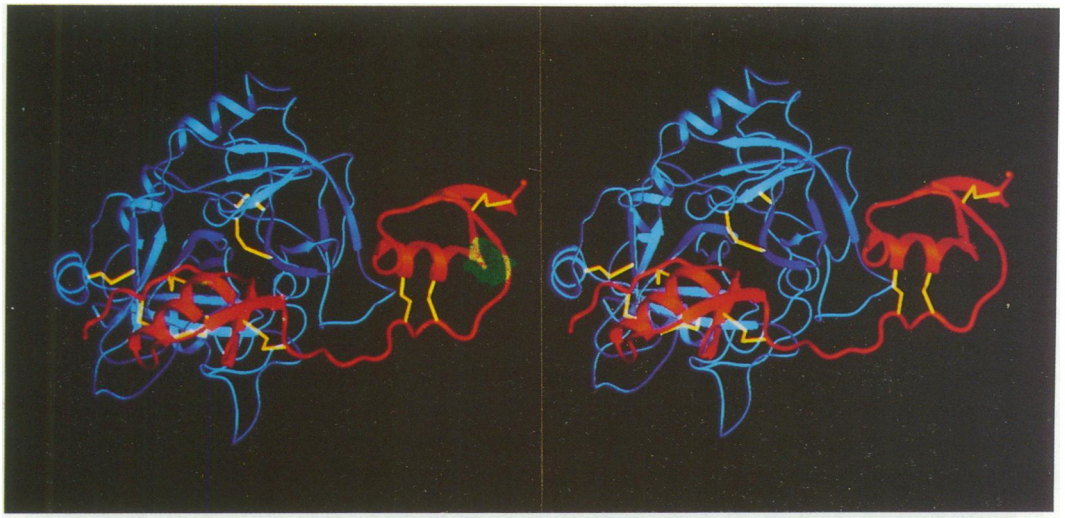

Fig. 1. Stereo view of the complex formed between thrombin (blue) and rhodniin (red) in the thrombin standard orientation, i.e. with the active-site cleft facing the viewer and a bound inhibitor chain running from left to right. Yellow connections indicate disulfide bridges. Rhodniin interacts through its amino-terminal domain in a canonical manner with the active site and through its carboxy-terminal domain with the fibrinogen recognition exosite of thrombin. The secondary structure elements were determined with the program DSSP (Kabsch and Sander, 1983). The figure was produced with SETOR (Evans, 1993).

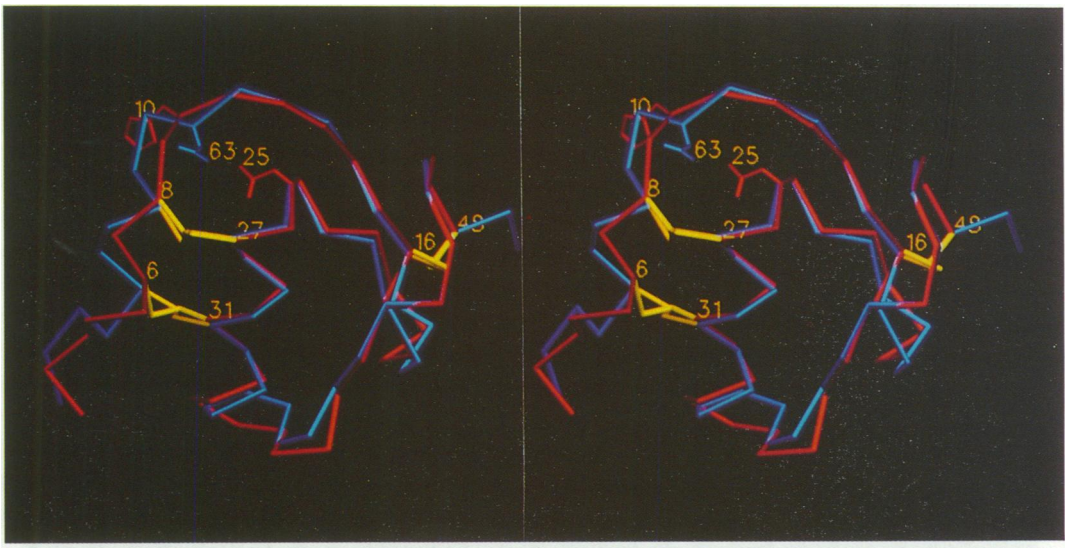

Fig. 2. Superimposition of the amino-terminal (red) and carboxy-terminal (blue) domain of rhodniin. Both domains are shown as C_α-models. The side chains of His10I (P1) and Asp63I are also drawn. The figure was produced using SETOR (Evans, 1993).

late oxygen of this Glu192 is hydrogen bonded to another buried water molecule, which is thus surrounded in a tetrahedral manner by four hydrogen bond partners, namely two hydrogen bond acceptors (Glu192 carboxylate oxygen and Glu146 O) and two hydrogen bond donors (Asn143 N^{δ} and Ser24I O^{γ}) (see Figure 3). Thus His10I(P1), in spite of the normally observed high preference of thrombin for Arg P1 residues, occupies the S1 pocket and simultaneously satisfies the special hydrogen bond and charge compensation requirements set by Glu192.

Similar to other related protein inhibitor–serine proteinase complexes (see Bode and Huber, 1992), the scissile peptide bond His10I–Ala11I of rhodniin seems to be intact in its thrombin complex. However, there is continuous electron density at the 2σ contouring level between Ser195 O^{γ} and His10I C, which might indicate a substantial degree of covalent interaction. This possibility is supported by protein refinement with relaxed geometry restraints, which results in an approach of the O^{γ} atom to 2 Å from this carbonyl carbon. Due to the restricted resolution of the reflection data, however, this statement must be considered with the appropriate caution.

The side chain of P4 residue Ala7I points towards thrombin's aryl-binding site lined by Ile174, Trp215 and segment 97–99; the small side chain thus leaves this site empty in the complex. In contrast to this non-complementary arrangement up to P4, the central and the primed part of the reactive-site loop of rhodniin's amino-terminal domain fit into the narrow active-site cleft (see Figure 4). Cys8I(P3) and its disulfide bridge to Cys27I directly pack against the indole moiety of Trp60D, shifting it slightly from the position observed in PPACK–thrombin; the pyrrolidine ring of Pro9I(P2) neatly fits into the S2 cavity of thrombin, formed by Leu99, His57, Tyr60A and Trp60D. The Ala11I(P1′) side chain fills the small S1′ pocket of thrombin left by Lys60F, and the His13I(P3′)

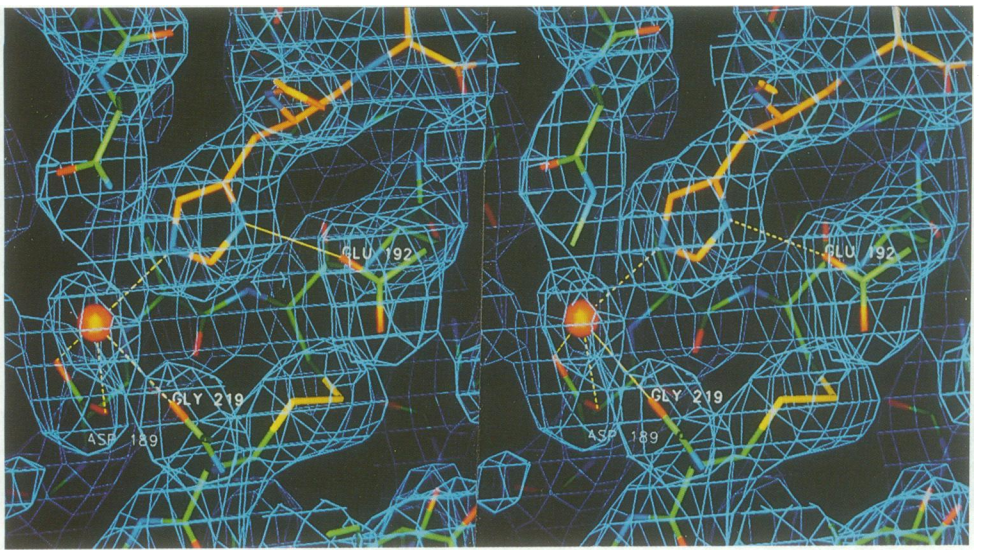

Fig. 3. The final coordinates of the S1 pocket of thrombin and its environment are shown with the $2F_o$–F_c density contoured at 1σ. Thrombin and rhodniin residues are displayed white and yellow respectively. Possible intermolecular hydrogen bonds are shown with broken lines. The imidazole group of His10I(P1) is hydrogen-bonded to the carboxylate group of Glu192 and to a water molecule, which fills the empty space in the pocket. This water molecule is also hydrogen-bonded to Asp189 at the bottom of the pocket and to the carbonyl group of Gly219. The figure was produced using the program MAIN (Turk, 1992).

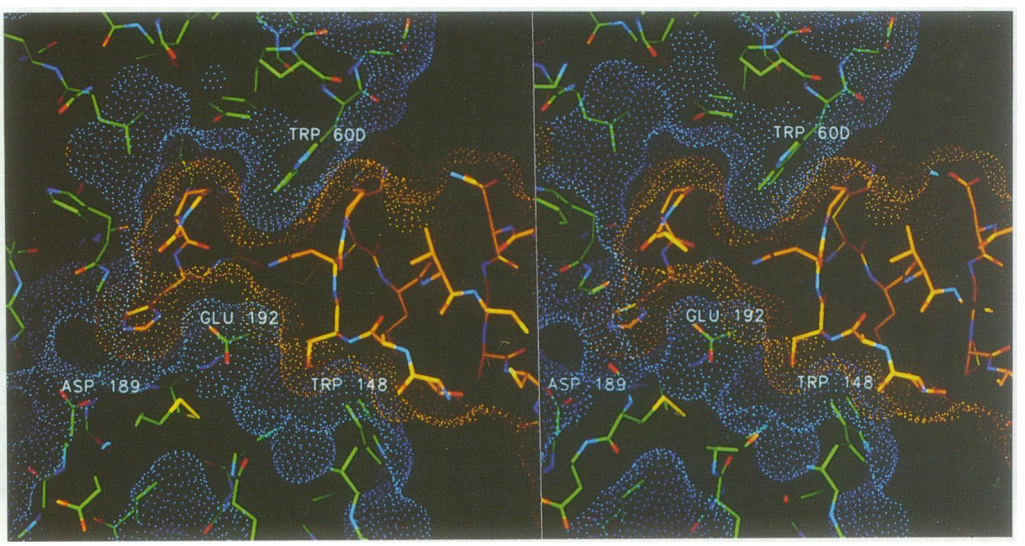

Fig. 4. Stereo view of a 10 Å section through the rhodniin (yellow)–thrombin (orange) complex, overlain with their corresponding Connolly dot surfaces. Thrombin's active-site residues, rhodniin's reactive site and Trp60D (top) and Trp148 (bottom) are shown. View is from the left side of the complex, i.e. 90° rotated along *y* compared with the standard view (Figure 1), so that the active-site cleft points to the right. The figure was produced with MAIN.

imidazole group can receive a hydrogen bond from the same lysine. The guanidine group of Arg14I(P4′) is favourably placed to form a hydrogen/ionic bond to thrombin's carboxylate group of Glu39, which in turn forms another hydrogen/ionic bond to the side chain of Arg35 flanking it on the opposite side.

In addition to this reactive-site loop, the first rhodniin domain is in direct contact with thrombin through residues Thr22I to Cys27I. In particular Pro26I forms, together with the His13I(P3′) imidazole moiety and Pro9I(P2), a groove which is interlocked with the indole group of Trp60D. The opposite flat side of the inhibitor domain around Thr22I touches the hydrophobic face of Trp148 and a few residues of the following 149 insertion loop (see Figure 4). This characteristic thrombin loop, which in PPACK–thrombin (Bode *et al.*, 1989; 1992) and in hirudin–thrombin (Grütter *et al.*, 1990; Rydel *et al.*, 1990, 1991) exhibits a narrow and quite open geometry respectively, adopts in the rhodniin complex a conformation intermediate between both extremes, presumably upon adapting to the inhibitor surface.

The carboxy-terminal domain of rhodniin

The carboxy-terminal domain of rhodniin is almost identical in size and shape to the amino-terminal domain. It likewise includes a core made up of a three-turn α-helix and a three-stranded antiparallel β-pleated sheet, with the N-terminal part of the helix enclosed by the curved reactive-site loop (see Figure 2). This loop segment differs, however, in that two additional single residues are

inserted; one between the first and the second cysteine, and the other behind this second cysteine (see Table I). Both insertions are accommodated, almost without distortion of the residual scaffold, by bulges resulting in a S-like curled loop segment Asp55I–Asp63I consisting of three adjacent 1–4 tight turns. As a consequence, this loop cannot bind to a serine proteinase. Both amino-terminal cysteine residues, Cys57I and Cys60I, and their side chains are nevertheless oriented to form similar disulfide bridges to the central helix as the equivalent residues do in the first domain, whereas the succeeding Asp61I–Gly62I–Asp63I segment differs considerably from Pro9I(P2) and His10I(P1) (see Figure 2). From Glu64I onwards, however, the loop segment runs virtually identically to the P1′–P4′ segment of the first domain. We note that sequence Glu64I–Tyr65I–Lys66I–Pro67I is almost identical to the P1′–P4′ segment of many ovomucoid third domains (Laskowski *et al.*, 1990). Loop segment Thr75I–Cys80I which crosses below the 'reactive-site loop' has a conformation similar to the equivalent loop in the first domain (see Figure 2); this holds also for the characteristic 'spacer' residue Asn78I, which hydrogen bonds via its carboxamide N^{δ} to Glu64I O (formally P1′), but finds only a water molecule where normally the P2-carbonyl is bound, due to the loop distortion caused by the insertions.

Interaction of the carboxy-terminal rhodniin domain with thrombin

The carboxy-terminal domain binds via its 'left' edge and 'left-hand front' side (relative to the orientation shown in Figure 2) to the positively charged patch extending to the right (thrombin standard orientation, see Figure 1) of the 37 loop, which constitutes most of the fibrinogen recognition exosite. There the amino-terminal segment up to Asp63I is placed on top of thrombin segment Arg75–Arg77A, while the surface residues of the third turn of the helix and the following loop are in contact with thrombin residues around Arg67 and segment 80–84 through a flat, contourless interface. The amino-terminal decapeptide including five acidic residues is placed across the positive charge cluster formed by thrombin residues Arg67, Arg75, Arg77A and Lys87; only Glu59I (2.8 Å to Arg75 N^{η}) and Asp63I (3.1 Å to Arg77A N^{η}) are involved in direct ionic/hydrogen bonds, while the side chains of Asp55I, Asp61I and Glu64I point away from the thrombin surface and form intramolecular salt bridges with Arg81I and Lys66I. Glu83I and Glu94I, anchored in the middle of the helix and in the following loop respectively, are oriented towards residues Lys81, and Lys109, Arg110, Lys36 at the opposing thrombin surface, but do not form direct charged hydrogen bonds with these residues. The four charged side chains of Lys96I, His97I, Glu98I and Arg102I at the carboxy terminus of the carboxy-terminal domain point away from the thrombin surface; their overall positive charge might assist pre-orientation of the second rhodniin domain with respect to the thrombin exosite I. It should be noted that Asp55I and Asp61I, in addition to forming the intramolecular salt bridge with Arg81I, make ionic/hydrogen bond contacts with Lys169 N^{ζ} of a symmetry-related thrombin molecule, and that a few atoms of the third helical turn and the carboxy-terminal residues are in crystal contacts with adjacent molecules; these contacts apparently assist in fixing the rhodniin second domain in the crystals, but seem to be too weak to affect overall binding of the carboxy-terminal domain to the thrombin exosite.

The connecting inter-domain linker peptide

The six residue linker peptide is spanned between the last residue of the first domain, Cys48I, and the first turn-forming residue of the second domain, Asp55I, with an overall extended conformation such that the anionic side chains alternately point in opposite directions and run approximately parallel to the thrombin surface, albeit at a significant distance. Although the rhodniin linker main chain follows a zig-zag course similar to hirudin segment 52I–57I in the corresponding hirudin–thrombin complexes, the rhodniin linker is shifted out of the active-site cleft and 'down' (seen in the thrombin standard orientation used in Figure 1) for ~5 Å, so that only the Asp51I side chain [which in addition forms an internal salt bridge with Arg14I(P4′)] directly touches thrombin surface residues, though only through a few van der Waals contacts. In a complex with human thrombin, where a lysine residue replaces the Glu149E of the bovine species, Glu49I could become engaged in a more direct interaction.

Discussion

A large number of multi-domain serine proteinase inhibitors is known, in particular of the Kazal type where up to seven domains can be loosely connected via more or less flexible linker peptides (see Laskowski *et al.*, 1981). Rhodniin represents the first protein inhibitor where two flexibly linked domains bind to the same proteinase module. Both domains belong to the Kazal family, but possess truncated amino-terminal segments compared with the classical Kazal inhibitors.

The amino-terminal domain of rhodniin fits precisely into the active-site cleft of thrombin. This is due not only to the thin binding edge characteristic of Kazal-type inhibitors (see Bode *et al.*, 1992), but results in particular from the specific sequence of distinct reactive-site loop residues. In addition to the interactions of Pro9I, His10I and Ala11I, the side chain of Leu12I fits the S2′ site better than the more common Tyr; His13I can mediate hydrogen bonds and stack with aromatic residues in S3′, and Arg14I at P4′ allows charge compensation of Glu39. The clustering of positively charged inhibitor residues at P3′ and P4′ might be particularly beneficial for thrombin binding; an aspartic acid in the P3′ position of the reactive-site loop markedly reduces the inhibitory potency of antithrombin III (Theunissen *et al.*, 1993), and mutational analysis of Glu39 to Lys in thrombin (Le Bonniec *et al.*, 1991) has suggested that this acidic residue might prefer a basic residue at P3′.

Besides the good complementarity of the reactive-site loop, the tightness of the rhodniin first domain–thrombin complex seems also to be achieved by the good fit of secondary binding sites with both exposed insertion loops, increasing the buried surface area to ~1200 $Å^2$, i.e. well above the values normally observed for canonical inhibitors (Bode and Huber, 1992).

In light of the strong preference of thrombin for P1 Arg residues in peptide substrates, the presence of a P1

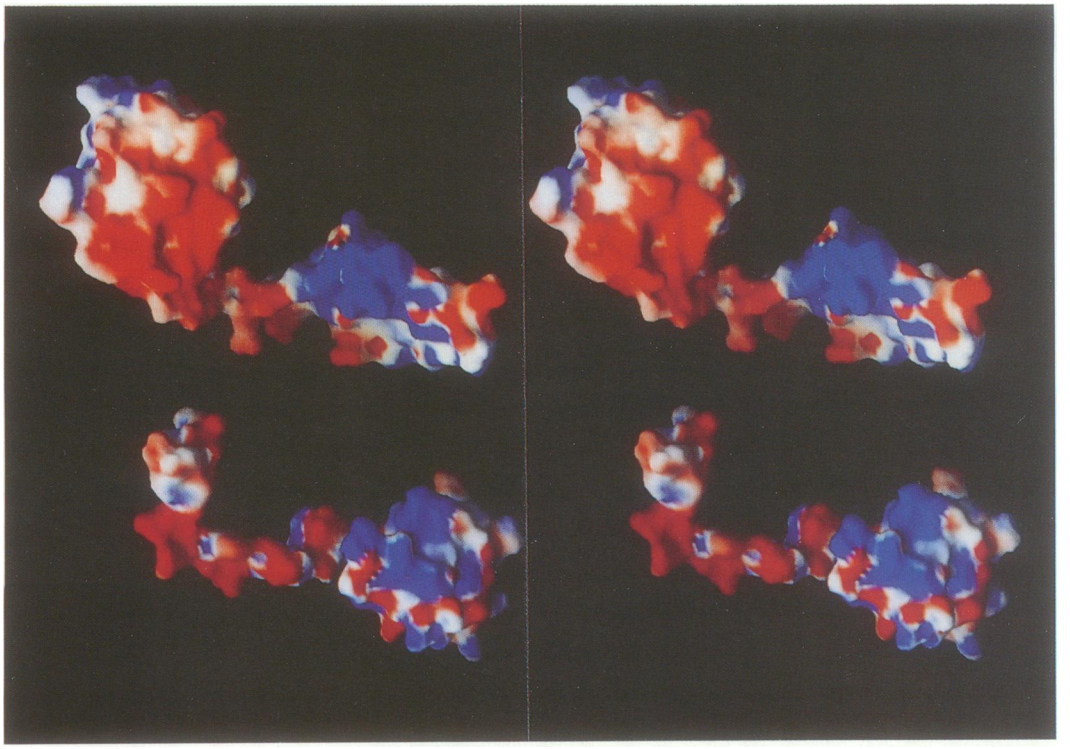

Fig. 5. Stereo plot of rhodniin (top) and hirudin (bottom) showing the thrombin-binding sides of both inhibitors by their solid surface. Both inhibitors are viewed at as seen by the thrombin molecule, i.e. rotated 180° around the vertical axis compared with Figure 1. The calculation and graphical representation were done with the program Grasp (Nicholls *et al.*, 1993) using the Amber partial charges (Weiner *et al.*, 1984), dielectric constants of 2 and 80 for the protein and solvent respectively, and a salt concentration of 0.145 M considering solvent accessibility and non-accessibility for the inhibitors in their thrombin complexes. The electrostatic potentials were contoured from –8 (intense red) to 8 k_BT/e (intense blue), where k_B is Boltzmann's constant and T the absolute temperature. The contact surface of the amino-terminal rhodniin domain is lined by both negative and positive groups, whereas that of the carboxy-terminal domain is dominated by negative charges.

His residue in the first rhodniin domain is astonishing on first glance only. To our knowledge, His has never been found as a P1 residue in any of the naturally occurring serine–proteinase protein inhibitors (see Laskowski *et al.*, 1987), and the occurrence of a P1 His residue at the cleavage site of various deficient fibrinogen species is associated with impaired fibrin polymerization (Lane *et al.*, 1993), suggesting that the conservation of Arg arises from evolutionary constraints other than protease binding. The rhodniin–thrombin structure reveals that the histidine side chain can be suitably arranged in the S1 pocket, with a water molecule coupling its imidazole group to Asp189 at the bottom. A very similar arrangement had previously been suggested for thrombin's interaction with Arg16→His dys-fibrinogenaemic fibrinogen variants by modelling on the basis of the trypsin–pancreatic trypsin inhibitor complex (Southan *et al.*, 1985). It might be interesting to note that inhibitors with clear homology to rhodniin have been isolated from *Triatoma infestans*, *Triatoma megista* and *Triatoma phyllosoma*, which display an arginine at the P1 position, and which, upon isolation via a thrombin affinity column, were partly found with the reactive-site bond cleaved, in strong contrast to rhodniin. This observation might indicate the importance of a special P1 side chain-S1 pocket arrangement for catalysis, capable of pulling the arginine side chain into the pocket with simultaneous exertion of distortion forces on the scissile peptide bond.

The particular highlight of the rhodniin–thrombin structure is, however, the arrangement of the His10I imidazole group near thrombin's characteristic Glu192 side chain; in this way meeting all hydrogen bonding and charge compensation requirements. The favourable arrangement of the Glu192 carboxylate group in this complex is assisted by an additional buried solvent molecule, which connects it to other groups in the vicinity. The dramatically changed properties of a Glu192→Gln thrombin mutant towards protein C and factor X activation clearly showed that Glu192 is a major determinant for restricting thrombin's specificity (Le Bonniec and Esmon, 1991). Thrombin apparently shares this acidic residue only with protein C, whereas the majority of trypsin-like enzymes carry a glutamine at this position. In bovine pancreatic trypsin inhibitor complexes with trypsin, the equivalent Gln192 side chain hydrogen bonds via its amide nitrogen to the P2 carbonyl oxygen and compensates the parallel dipole moments of this and of adjacent carbonyl groups (Huber *et al.*, 1974; Bode *et al.*, 1984). The Glu192 carboxylate side chain of wild-type thrombin cannot form such a hydrogen bond; in fact, in PPACK–thrombin the Glu192 side chain does not interact with the PPACK inhibitor, but extends freely into solution (Bode *et al.*, 1989, 1992). Consistent with this role, the Glu192→Gln mutant binds the bovine pancreatic trypsin inhibitor several orders of magnitude more tightly than the wild-type enzyme (Guinto *et al.*, 1994).

Altogether, the binding of the first rhodniin domain to the thrombin active site seems to be governed by shape complementarity, favourable hydrogen bonding and electrostatic interactions with opposite polarities in the non-primed and primed regions (see Figure 5).

The carboxy-terminal domain interacts with the fibrinogen recognition exosite of thrombin via relatively contourless faces. Of the many potential salt bridges in this rhodniin–thrombin arrangement, only two are directly made between the proteins, which together with a few intermolecular hydrogen bonds might provide directionality determining their arrangement. Preliminary electrostatic potential calculations indicate that most of these negative charges still feel the strong positive potential created by the thrombin exosite, and vice versa, and that this electrostatic contribution might be even more important for rhodniin than for hirudin, due to the more dense accummulation of negative charges on the interacting 'inner' rhodniin face compared with that of the hirudin tail (see Figure 5). This overall dominance of the electrostatic interactions in binding of the carboxy-terminal rhodniin domain is also suggested by the quite strong ionic strength dependence of the complex formation (J.Dodt, personal communication).

Quite peculiar is the arrangement of the extremely acidic linker peptide, which was expected to align along the exosite furrow running around the exposed 37 loop of thrombin, as observed for the tail segment Asp55–Ile59 of hirudin or the exosite binding pentapeptide Lys51′–Phe55′ of the thrombin receptor (Qiu *et al.*, 1992; Mathews *et al.*, 1994). The crystal structure of the rhodniin–thrombin complex instead reveals that this linker is kept at a considerable distance from the exosite furrow. Most of the acidic side chains extend parallel to the thrombin surface instead of toward it, but their negative charges will certainly contribute to overall binding. An evolutionary advantage of this design might be that the linker segment does not have to be adapted to the restrictive exosite consensus sequence, characterized by several positively charged residues interspersed with hydrophobic crevices (see Stubbs and Bode, 1993).

Due to charge–charge repulsion, this highly negatively charged inter-domain linker might rather act as a spacer element that separates the domains in the non-complexed state into an extended conformation ready for thrombin binding. Similar to hirudin (Stone and Hofsteenge, 1986), rhodniin and thrombin might also mutually orient one another upon approach through electrostatic interactions. The second domain might first bind to the fibrinogen recognition exosite in a less restricted manner, before the amino-terminal domain would fit into thrombin's active-site cleft. With the individual domains binding in the nanomolar to micromolar range (data not shown), the additive contributions would result in the tight-binding rhodniin–thrombin complex. Further kinetic experiments with both independent domains and crystallographic studies are needed to clarify the detailed association steps.

Materials and methods

Bovine thrombin was prepared from ox blood as described by Brandstetter *et al.* (1992). Recombinant rhodniin was expressed in the yeast *Hansenula polymorpha* (Rhein Biotech GmbH, Düsseldorf) and purified according to Friedrich *et al.* (1993). Thrombin was co-crystallized with a slight excess of a 1:1 molar ratio of rhodniin to thrombin. Orthorhombic crystals (space group $P2_12_12_1$, two complexes per asymmetric unit) and monoclinic crystals (space group C2, two complexes per asymmetric unit) were grown at 20°C as needles from 10% PEG 4000, 300 mM phosphate buffer pH 6, in hanging drops using the vapour diffusion technique. Diffraction data up to 3.1 Å for the orthorhombic crystals and 2.6 Å for the monoclinic ones were collected on a MAR imaging plate system, evaluated using the Mosflm package (Leslie, 1994) and ROTAVATA from the CCP4 Suite (CCP4, 1994), and loaded and scaled with PROTEIN (Steigemann, 1974).

Table II. Crystal data and refinement parameters for the rhodniin–thrombin complexes

Spacegroup	C2	$P2_12_12_1$
Cell constants (Å)		
a	114.74	91.311
b	112.3	111.6
c	92.07	112.14
β (°)	94.97	
Limiting resolution (Å)	2.6	3.1
Significant measurements	89438	31602
R_{merge}[a] (%)	4.9	6.2
Independent reflections	33844	16957
Completeness	94.5% (∞–2.6 Å)	79.1% (∞–3.1 Å)
outermost shell	78.2% (2.66–2.6 Å)	82.3% (3.17–3.1 Å)
Non-hydrogen protein atoms	6530	6530
Solvent molecules	203	73
Reflections used for refinement	30087	15063
Resolution range (Å)	8.0–2.64	8.0–3.1
Completeness (%)	91.3	75.1
R value[b] (%)	18.9	17.5
R_{free} (%)	26.2	
R.m.s. standard deviation		
bond lengths (Å)	0.010	0.010
bond angles (°)	1.66	1.62
R.m.s.B[c] (Å^2)	3.83	1.96

[a] $R_{merge} = \Sigma[I-\langle I\rangle]/\Sigma I$.
[b] R value $= \Sigma(|F_{obs}|-|F_{calc}|)/\Sigma|F_{obs}|$.
[c] R.m.s.B, r.m.s. deviation of the B-factor of bonded atoms.

The structure was solved by Patterson search techniques. Rotational and translational searches for the orientation and position of the thrombin molecules in the monoclinic crystals were performed with the program AMoRe (Navaza, 1994) using data up to 3.5 Å and the bovine thrombin model as previously obtained in complex with NAPAP (Brandstetter *et al.*, 1992). The rotational search showed two solutions with correlation values of 0.22 and 0.20 over 0.09 for the next highest peak. Translational search and rigid body fitting for these two solutions resulted in a correlation value of 0.54, with the two independent complex molecules related by a 176° rotation axis almost parallel to *c*. The initial phases obtained from the thrombin molecules were good enough to trace (in addition to the two thrombin molecules) parts of the amino-terminal domain of rhodniin closely bound to thrombin. An envelope covering thrombin and rhodniin's amino-terminal domain and the region where the carboxy-terminal domain was expected to be was created, and the density between the two molecules in the asymmetric unit was averaged using the programs MAIN (Turk, 1992) and PROTEIN. Several refinement cycles, consisting of model-building using O (Jones *et al.*, 1991) and conjugate gradient minimization with X-PLOR (Brünger, 1992a) using the parameters of Engh and Huber (1991), reduced the *R* value to 0.189. In the later stages restrained individual B-factor refinement was applied. The free *R* value (Brünger, 1992b) was used with 10% of the reflections to gauge the progress of the initial refinement steps. The final free *R* value was calculated as 0.262. Two hundred and three water molecules were placed at stereochemically reasonable positions.

For the orthorhombic data, the rotational and translational search was also performed with AMoRe using the same model as in the monoclinic case. Only one rotational solution was found with a correlation value of 0.22 over 0.12 for the next highest peak. The translational search yielded two solutions (correlation values of 0.26 and 0.37), with the two molecules having almost the same orientation, but shifted by almost 1/2 along *b* and *c*, simulating a pseudo centered lattice. Seventy-three water molecules from the monoclinic model were included in the orthorhombic models that showed peaks of at least 2σ in the difference density. The orthorhombic data were refined to a final *R* value of 0.175 using X-PLOR (see Table II).

In one of the two rhodniin copies present in the monoclinic crystals, the complete rhodniin chain can be traced in the electron density, whereas there is no density for the first three amino acids in the second copy.

The rhodniin structure is virtually the same in all four copies examined here. Therefore, the description of structural features will refer to the complex with the better defined rhodniin molecule (molecule 2 in the monoclinic crystal form).

The coordinates of both crystal forms have been deposited with the Brookhaven Protein Data Bank.

Acknowledgements

We thank U.Baumann, H.Brandstetter, R.Engh, D.Turk and especially M.Stubbs for stimulating and fruitful discussions and R.Engh, T.Mather and M.Stubbs for critical reading of this manuscript. This work has been supported by the Sonderforschungsbereich 207. The financial support (to D.L.) provided by a NATO–CNR Junior fellowship is gratefully acknowledged.

References

Barton,G.J. (1993) ALSCRIPT a tool to format multiple sequence alignments. *Protein Engng*, **6**, 37–40.

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

Bode,W., Walter,J., Huber,R., Wenzel,H. and Tschesche,H. (1984) The refined 2.2 Å X-ray structure of the ternary complex formed by bovine trypsinogen, valine–valine and the Arg15 analogue of bovine pancreatic trypsin inhibitor. *Eur. J. Biochem.*, **144**, 185–190.

Bode,W., Mayr,I., Baumann,U., Huber,R., Stone,S.R. and Hofsteenge,J. (1989) 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., Turk,D. and Karshikov,A. (1992) The refined 1.9 Å X-ray crystal structure of D-PheProArg chloromethylketone inhibited human α-thrombin. Structure analysis, overall structure, electrostatic properties, detailed active site geometry, structure-function relationships. *Protein Sci.*, **1**, 426–471.

Bolognesi,M., Gatti,G., Menegatti,E., Guarneri,M., Marquart,M., Papamokos,E. and Huber,R. (1982) Three-dimensional structure of the complex between pancreatic secretory trypsin inhibitor (Kazal-type) and Trypsinogen at 1.8 Å resolution. Structure solution, crystallographic refinement and preliminary interpretation. *J. Mol. Biol.*, **162**, 839–868.

Brandstetter,H., Turk,D., Höffken,H.W., Grosse,D., Stürzebecher,J., Martin,P.D., Edwards,B.F.P. and Bode,W. (1992) Refined 2.3 Å X-ray crystal structure of bovine thrombin complexes formed with the benzamidine and arginine-based thrombin inhibitors NAPAP, 4-TAPAP and MQPA. *J. Mol. Biol.*, **226**, 1085–1099.

Brünger,A.T. (1992a) X-PLOR, *Version 3.1. A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.

Brünger,A.T. (1992b) Free *R* value: a novel statistic quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–474.

The CCP4 Suite: Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr.*, **D50**, 760–763.

Church,F.C., Pratt,C.W., Neyes,C.M., Kalayanamit,T., Sherrill,G.B., Tabin,R. and Meade,J.B. (1989) Structural and functional properties of human α-thrombin, and γ_T-thrombin. *J. Biol. Chem.*, **264**, 18419–18425.

Dodt,J., Köhler,S. and Baici,A. (1988) Interaction of site specific hirudin variants with α-thrombin. *FEBS Lett.*, **229**, 87–90.

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

Evans,S.V. (1993) SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graphics*, **11**, 134–138.

Friedrich,T., Kröger,B., Bialojan,S., Lemaire,H.G., Höffken,H.W., Reuschenbach,P., Otte,M. and Dodt,J. (1993) A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. *J. Biol. Chem.*, **268**, 16216–16222.

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

Guinto,E.R., Ye,J., Le Bonniec,B.F. and Esmon,C.T. (1994) Glu192 → Gln substitution in thrombin yields an enzyme that is effectively inhibited by bovine pancreatic trypsin-inhibitor and tissue factor pathway inhibitor. *J. Biol. Chem.*, **269**, 18395–18400.

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

Huber,R., Kukla,D., Bode,W., Schwager,P., Bartels,K., Deisenhofer,J. and 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.

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

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

Kalafatis,M., Swords,N.A., Rand,M.D. and Mann,K.G. (1994) Membrane-dependent reactions in blood-coagulation – role of the vitamin-K-dependent enzyme complexes. *Biochim. Biophys. Acta*, **1227**, 113–129.

Kato,I., Kohr,W.J. and Laskowski,M.,Jr (1978) In Magnusson,S., Ottesen,M., Foltman,B., Dano,K. and Neurath,H. (eds), *Regulatory Proteolytic Enzymes and Their Inhibitors*. 11th FEBS Meeting. Pergamon Press, Oxford, Vol. XLVII, pp. 197–206.

Kazal,L.A., Spicer,D.S. and Brahinsky,R.A. (1948) Isolation of a crystalline trypsin inhibitor – anticoagulant protein from pancreas. *J. Am. Chem. Soc.*, **70**, 3034–3040.

Lane,D.A. *et al.* (1993) Antithrombin III mutation database: First update. For the thrombin and its inhibitors subcommittee of the scientific and standardization committee of the International Society on Thrombosis and Haemostasis. *Thromb. Haemostasis*, **70**, 361–369.

Laskowski,M.,Jr *et al.* (1981) Correlation of amino acid sequence with inhibitor activity and specificity of protein inhibitors of serine proteinases. In Eggerer,H. and Huber,R. (eds), *Structural and Functional Aspects of Enzyme Catalysis*. Colloquium, Mosbach, Vol. XXXII, pp. 136–152.

Laskowski,M.,Jr *et al.* (1990) Amino acid sequences of ovomucoid third domain from 25 additional species of birds. *J. Protein Chem.*, **9**, 715–725.

Le Bonniec,B.F. and Esmon,C.T. (1991) Glu192 → 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. and Esmon,C.T. (1991) Thrombin Glu39 restricts the P'3 specificity to nonacidic residues. *J. Biol. Chem.*, **266**, 13796–13803.

Leslie,A.G.W. (1994) *Mosflm User Guide, Mosflm version 5.20*. MRC Laboratory of Molecular Biology, Cambridge, UK.

Lottenberg,R., Hall,J.A., Blinder,M., Binder,E.P. and 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.

Markwardt,F. (1994) Coagulation inhibitors from blood-sucking animals. *Pharmazie*, **49**, 313–316.

Mathews,I.I. *et al.* (1994) Crystallographic structures of thrombin complexed with thrombin receptor peptides: existence of expected and novel binding modes. *Biochemistry*, **33**, 3266–3279.

Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.

Nicholls,A., Bharadwaj,R. and Honig,B. (1993) Grasp – graphical representation and analysis of surface properties. *Biophys. J.*, **64**, A166.

Qiu,X., Padmanabhan,K.P., Carperos,V.E., Tulinsky,A., Kline,T., Maraganore,J.M. and Fenton,J.W.,II (1992) Structure of the hirulog 3-thrombin complex and nature of the S' subsites of substrates and inhibitors. *Biochemistry*, **31**, 11689–11697.

Rydel,T.J., Ravichandran,K.G., Tulinsky,A., Bode.W., Huber,R., Roitsch,C. and 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. and Huber,R. (1991) Refined structure of the hirudin–thrombin complex. *J. Mol. Biol.*, **221**, 583–601.

Scacheri,E., Nitti,G., Valsasina,B., Orsini,G., Visco,C., Ferrera,M., Sawyer,R.T. and Sarmientos,P. (1993) Novel hirudin variants from the leech *Hirudinaria manillensis* – amino-acid-sequence, cDNA cloning and genomic organization. *Eur. J. Biochem.*, **214**, 295–304.

Southan,C., Lane,D.A., Bode,W. and Henschen,A. (1985) Thrombin-induced fibrinopeptide release from a fibrinogen variant (fibrinogen Sydney I) with an AxArg-16 → His substitution. *Eur. J. Biochem.*, **147**, 593–600.

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, Technische Universität, München.

Steiner,V., Knecht,R., Bornsen,K.O., Gassmann,E., Stone,S.R., Raschdorf,F., Schlaeppi,J.M. and Maschler,R. (1992) Primary structure

and function of novel O-glycosylated hirudins from the leech *Hirudinaria manillensis*. *Biochemistry*, **31**, 2294–2298.

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

Strube,K.-H., Kröger,B., Bialojan,S., Otte,M. and Dodt,J. (1993) Isolation, sequence analysis, and cloning of haemadin. *J. Biol. Chem.*, **268**, 8590–8595.

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

Theunissen,H.J.M., Dijkema,R., Grootenhuis,P.D.J., Swinkels,J.C., de Porter,T.L., Carati,P. and Visser,A. (1993) Dissociation of heparin-dependent thrombin and factor Xa inhibitory activities of antithrombin III by mutations in the reactive site. *J. Biol. Chem.*, **268**, 9035–9040.

Turk,D. (1992) Weiterentwicklung eines Programms für Molekülgraphik und Elektronendichte-Manipulation und seine Anwendung auf verschiedene Protein-Strukturaufklärungen. Ph.D. Thesis, Technische Universität, München.

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

Walsmann,P. and Markwardt,F. (1981) Biochemische und pharmakologische aspekte des thrombininhibitors hirudin. *Pharmazie*, **36**, 653–660.

Weber,E., Papamokos,E., Bode,W., Huber,R., Kato,I. and Laskowski, M.,Jr (1981) Crystallization, crystal structure analysis and molecular model of the third domain of Japanese Quail Ovomucoid, a Kazal-type inhibitor. *J. Mol. Biol.*, **149**, 109–123.

Weiner,S., Kollman,P., Case,D., Singh,U., Ghio,C., Alagona,G., Profeta,S. and Weiner,P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, **106**, 765–784.

Ye,J., Rezaie,A.R. and Esmon,C.T. (1994) Glycosaminoglycan contributions to both protein C activation and thrombin inhibition involve a common arginine-rich site in thrombin that includes residues arginine93, arginine97 and arginine101. *J. Biol. Chem.*, **269**, 17965–17970.

Zingali,R.B., Jandrotperrus,M., Guillin,M.C. and Bon,C. (1993) Bothrojaracin, a new thrombin inhibitor isolated from *Botrops jararaca* venom – characterization and mechanism of thrombin inhibition. *Biochemistry*, **32**, 10794–10802.

Received on June 29, 1995; revised on July 28, 1995