A new structural class of serine protease inhibitors revealed by the structure of the hirustasin–kallikrein complex

Peer RE Mittl[1], Stefania Di Marco[1], Gabriele Fendrich[1], Gabriele Pohlig[1], Jutta Heim[1], Christian Sommerhoff[2], Hans Fritz[2], John P Priestle[1] and Markus G Grütter[1]*

Background: Hirustasin belongs to a class of serine protease inhibitors characterized by a well conserved pattern of cysteine residues. Unlike the closely related inhibitors, antistasin/ghilanten and guamerin, which are selective for coagulation factor Xa or neutrophil elastase, hirustasin binds specifically to tissue kallikrein. The conservation of the pattern of cysteine residues and the significant sequence homology suggest that these related inhibitors possess a similar three-dimensional structure to hirustasin.

Results: The crystal structure of the complex between tissue kallikrein and hirustasin was analyzed at 2.4 Å resolution. Hirustasin folds into a brick-like structure that is dominated by five disulfide bridges and is sparse in secondary structural elements. The cysteine residues are connected in an *abab cdecde* pattern that causes the polypeptide chain to fold into two similar motifs. As a hydrophobic core is absent from hirustasin the disulfide bridges maintain the tertiary structure and present the primary binding loop to the active site of the protease. The general structural topography and disulfide connectivity of hirustasin has not previously been described.

Conclusion: The crystal structure of the kallikrein–hirustasin complex reveals that hirustasin differs from other serine protease inhibitors in its conformation and its disulfide bond connectivity, making it the prototype for a new class of inhibitor. The disulfide pattern shows that the structure consists of two domains, but only the C-terminal domain interacts with the protease. The disulfide pattern of the N-terminal domain is related to the pattern found in other proteins. Kallikrein recognizes hirustasin by the formation of an antiparallel β sheet between the protease and the inhibitor. The P1 arginine binds in a deep negatively charged pocket of the enzyme. An additional pocket at the periphery of the active site accommodates the sidechain of the P4 valine.

Addresses: [1]CDDT, Pharmaceutical Research, Ciba-Geigy Ltd., 4002 Basel, Switzerland and [2]Abteilung für Klinische Chemie und Klinische Biochemie, Ludwig Maximilians Universität, Nussbaumstrasse 20, 80336 München, Germany.

*Corresponding author.
E-mail: gruetter@fmi.ch

Dedicated to R Huber on the occasion of his 60th birthday

Key words: crystal structure, disulfide pattern, hirustasin, PSA, serine protease inhibitor, tissue kallikrein

Received: 21 October 1996
Revisions requested: 19 November 1996
Revisions received: 5 December 1996
Accepted: 6 December 1996

Electronic identifier: 0969-2126-005-00253

Structure 15 Month 1997, 5:253–264

Introduction

The saliva of leeches contains a variety of protease inhibitors. In order to maintain the liquid state of the ingested blood, most of these inhibitors are directed against proteases that are involved in blood coagulation; for example, hirudin specifically inhibits thrombin [1], and decorsin inhibits platelet aggregation by blocking fibrinogen binding to the platelet receptor glycoprotein IIb-IIIa [2,3].

Recently, a 55 amino acid inhibitor, called hirustasin, has been isolated from the leech *Hirudo medicinalis* [4]. It is able to inhibit tissue kallikrein but not plasma kallikrein or plasmin. Hirustasin is the first inhibitor of tissue kallikrein that does not also inhibit plasma kallikrein. It has 27% and 32% sequence identity with the first and second domains, respectively, of the 119 amino acid inhibitor antistasin, which was originally isolated from the salivary glands of the Mexican leech *Haementeria officinalis* [5]. Antistasin contains 10 disulfide bridges [6] and its primary structure shows a twofold internal repeat, indicating that it has evolved by a gene duplication event. Antistasin is a competitive, high-affinity inhibitor of the serine protease coagulation factor Xa, and is slowly cleaved at a single position during the course of inhibition. A protein that is almost identical to antistasin has been isolated from the giant Amazonian leech *Haementeria ghilianii* and named ghilanten [7]. Hirustasin differs from both antistasin and ghilanten in that it has no inhibitory effect on factor Xa. Instead, it inhibits tissue kallikrein, trypsin, chymotrypsin and neutrophil cathepsin G. Hirustasin exhibits no apparent inhibitory activity against elastases [4], but shares 51% sequence identity with the human leukocyte and porcine pancreatic elastase-specific inhibitor guamerin [8]; this 57 amino acid inhibitor has been isolated recently from the Korean leech, *Hirudo nipponia*.

The inhibitor class comprising hirustasin, antistasin, ghilanten and guamerin is characterized by the exact conservation of the spacing of 10 cysteine residues within the sequence of the protein [8]. This homology suggests the existence of a common ancestor which diverged by gene duplication and mutation events to produce a family of inhibitors with different specificities: hirustasin, primarily for tissue kallikrein; antistasin/ghilanten, specifically for factor Xa; and guamerin, for elastase. Unlike the anticoagulant antistasin, which, like hirudin, is essential to maintain the liquid state of ingested blood inside the leech [9], the biological functions of guamerin and hirustasin are unknown.

Tissue kallikreins form a group of closely related serine proteases that exhibit a narrow range of substrate specificities [10]. They belong to the peptidase family S1, also known as the trypsin family, and to the glandular kallikrein subfamily [11]. Tissue kallikreins catalyze the liberation of a biologically highly active kinin, kallidin or lysyl-bradykinin, from ubiquitously occurring kininogens by cleavage of a Met–Lys and Arg–Ser bond [12]. The tissue kallikrein–kinin system contributes to the maintenance of normal blood pressure and, using transgenic mouse models, a direct link has recently been shown to exist between kallikrein gene expression and changes in blood pressure [13]. Porcine pancreatic kallikrein is by far the best studied of the glandular kallikreins [14], and the crystal structure of this protease, both alone and complexed with aprotinin, another kallikrein inhibitor, has been analyzed [15,16].

Here, we present the crystal structure of the complex between hirustasin and porcine tissue kallikrein refined at 2.4 Å resolution. This is the first three-dimensional (3D) structure of a member of the inhibitor class to which hirustasin belongs, and the first time that the interactions of a member of this class with the cognate protease have been determined.

Results and discussion

Overall hirustasin topology

The two hirustasin–kallikrein complexes present in the asymmetric unit (kallikrein 1, residues A16–B246; hirustasin 1, residues I5–I52; kallikrein 2, residues X16–Y246; hirustasin 2, residues J5–J52) have nearly identical conformations. A pairwise comparison of the Cα atoms of the protease and the inhibitor reveals root mean square (rms) deviations values of 0.24 Å and 0.36 Å, respectively. The larger rms deviation for the inhibitor is attributed to the weak electron density of the N terminus of hirustasin 2. Since this region is not stabilized by crystal packing interactions, it is highly flexible. The N terminus of hirustasin 1 shows much stronger electron density, but nevertheless possesses high B-factors. Except for the N terminus of hirustasin 2, both complexes are well defined in the electron-density map. As the Cα rms deviation is in the range of the estimated coordinate error, we consider the two complexes to be identical. Figure 1 shows an omit-map of the primary binding loop in the first complex. The following discussion focuses on the first complex (residues A16–B246 and I5–I52), because it possesses lower thermal mobility.

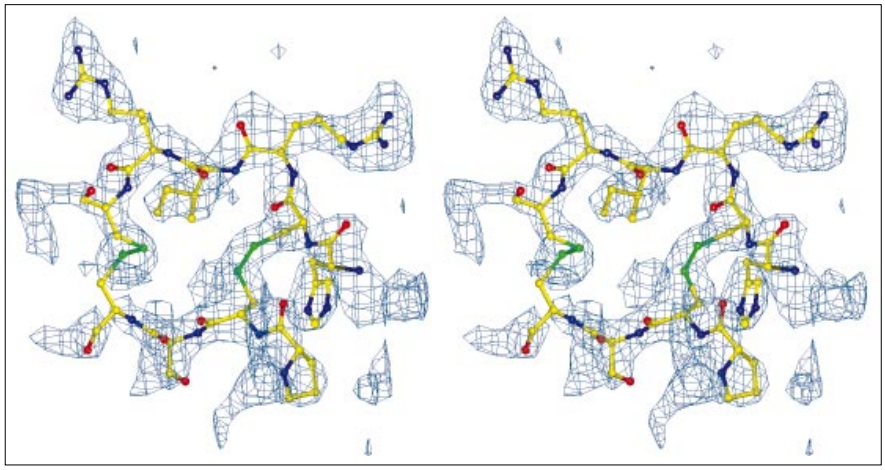
Hirustasin is a monomeric serine protease inhibitor that consists of 55 amino acid residues. In the crystal structure, four N-terminal and three C-terminal residues are not defined. The 48 residues that are seen in the electron-density map fold into a brick-shaped structure with overall dimensions of about $40\times20\times10$ Å^3. There are 10 cysteine residues involved in the formation of disulfide bridges (Fig. 2a,b). The hirustasin structure is relatively compact, but nevertheless can be subdivided into two domains. The N-terminal domain (residues 5–23) contains disulfide bridges Cys6–Cys17 and Cys11–Cys22, and the C-terminal domain (residues 23–52) contains disulfide bridges Cys24–Cys44, Cys29–Cys48 and Cys33–Cys50. Although these domains are much smaller than the classical domains [17], evidence for this subdivision comes from the Cα-distance matrix (data not shown), the disulfide bonding pattern (Fig. 3), and the similarity with structurally related inhibitors (Fig. 4). The average distances for Cα atoms within residues 5–23 and 23–52 are 9 Å and 11 Å, respectively, but the average distance for Cα atoms coming from the different parts is 19 Å. As the average distance between atoms from different parts is nearly twice as large, these fragments can be considered as independent domains. Residues 5–23 and 24–52 form domain N and domain C, respectively.

According to the nomenclature of Benham and Jafri [18], the cysteine residues are connected to form an *abab cdecde* pattern. This pattern contains an internal repeat of two sub-patterns. The *abab* and *cdecde* sub-pattern force the connected chains to run parallel. The parallel arrangement of residues 6–12 and 17–23 from domain N is repeated by residues 24–33 and 44–50 from domain C. When the N- and C-termini of a continuous polypeptide are running parallel to each other, the resulting helical turn can be either left- or right-handed. Surprisingly, both of the twofolds in domains N and C are left-handed. The spacing between subsequent disulfide bridges within a domain is nearly identical (4–5 residues), but the number of residues in the gap between cysteine residues from the same disulfide bridge differs. In domain N, this gap is 11 residues long in both cases, whereas in domain C it varies between 17 and 20 residues.

The structure of hirustasin contains few secondary structural elements. Only antiparallel β sheets are found and no α helices. In domain N, residues 14–25 form a twisted antiparallel β sheet and residues 6–9 and 12–15 form additional β turns. In the domain C, residues 38–46 make a

Figure 1

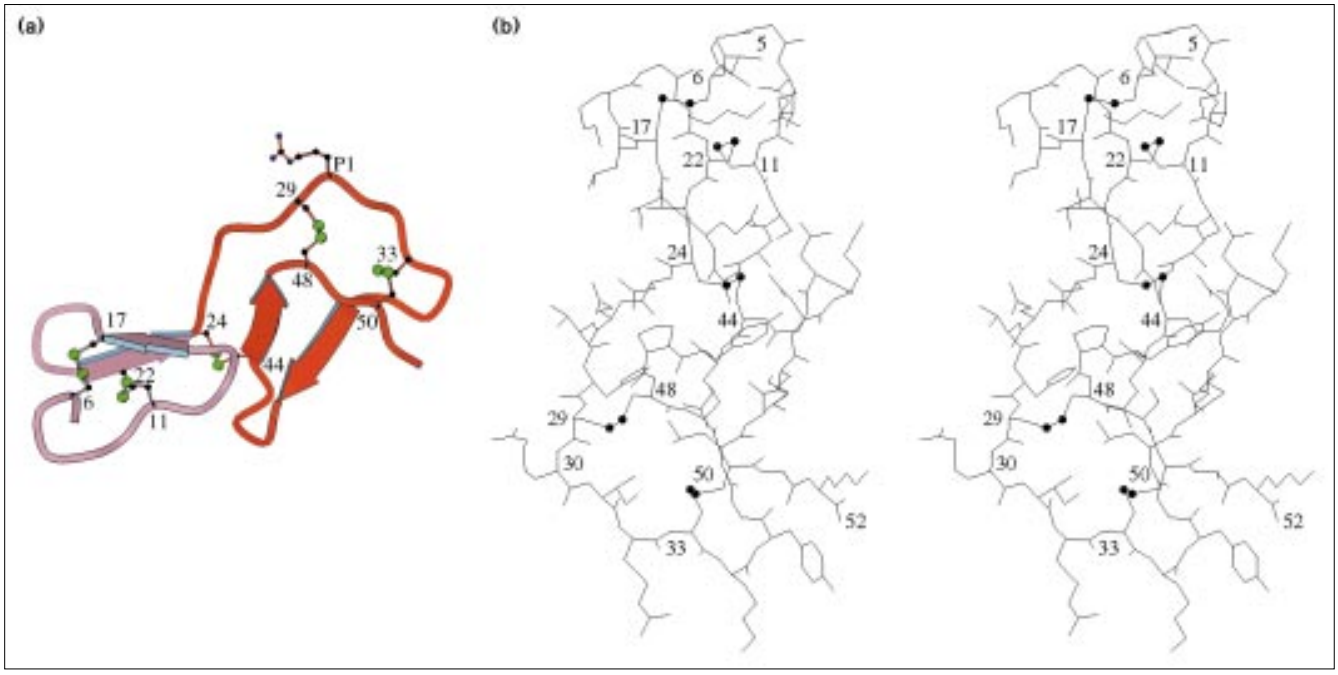
σ_A-weighted (F_o–F_c) omit map calculated in the resolution range 8.0–2.4 Å using calculated phases from the final model. The electron density within a 3 Å sphere around each atom from residues 28–33 and 47–50 is displayed at a contour-level of 2.5σ. Atoms are shown in standard colors.



short anti-parallel β sheet; residues 26–30 are also in a β sheet conformation, but the corresponding strand comes from the protease.

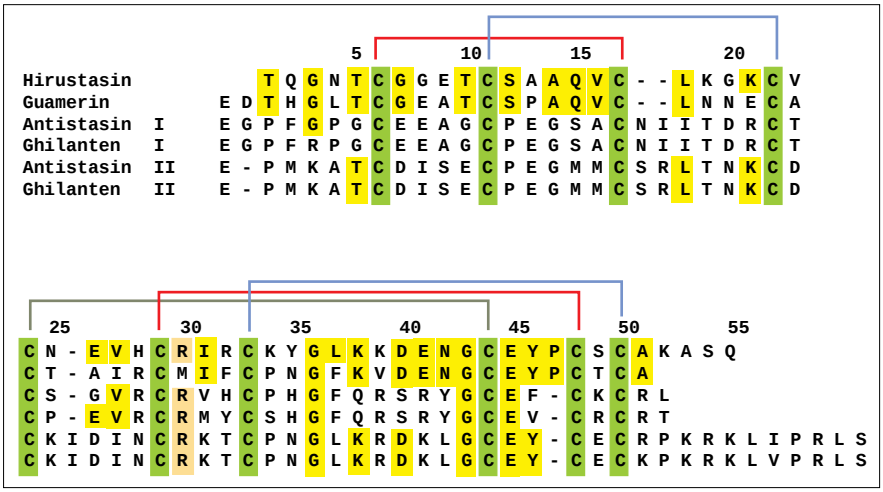
Although domains N and C possess similar disulfide patterns and secondary structure topologies, the structures are quite different because of different arrangements of disulfide bridges. The N-terminal part is described as a Cys_1-$loop_1$-Cys_2-$loop_2$-Cys_3-$loop_3$-Cys_4 motif with different lengths of the three loops. Disulfide bridges Cys6–Cys17 and Cys11–Cys22 form a central core which is decorated by loops. As in many other protease inhibitors [19], the hydrophobic core normally found in globular proteins is absent. The rigidity and the maintenance of a defined 3D structure is achieved by the disulfide bridges. Nevertheless, hydrophobic interactions play a role in the proper

Figure 2



The overall structure of hirustasin. (a) Ribbon diagram of hirustasin. The N-terminal and C-terminal parts are shown in different colors. The atom labels for the five disulfide bridges are given. The P1 residue Arg30 is indicated in the top right corner. (b) Stereo plot of the entire hirustasin chain. Numbers refer to cysteine residues and sulfur atoms are marked by bullets. The figure was generated using the program MOLSCRIPT [46].

Figure 3



Sequence alignment of hirustasin, guamerin and the two domains of antistasin and ghilanten (labeled I and II). Conserved residues, the P1 residue and the 10 cysteines are emphasized by yellow, brown and green bars, respectively. At the top, the connectivities of the disulfide bridges as they are observed in hirustasin are indicated.

folding of domain C. The disulfide bridges Cys29–Cys48 and Cys33–Cys50 from domain C form a ring-like structure, and the space inside the ring is filled by the side chains of Ile31 and Leu37. The ring also contains the primary binding loop that is present in many other serine protease inhibitors [20]. The disulfide bridge Cys24/Cys44 links the β sheet that flanks the ring to the N-terminal extension of the primary binding loop. Hydrogen bonds between mainchain atoms are restricted to the secondary structural elements.

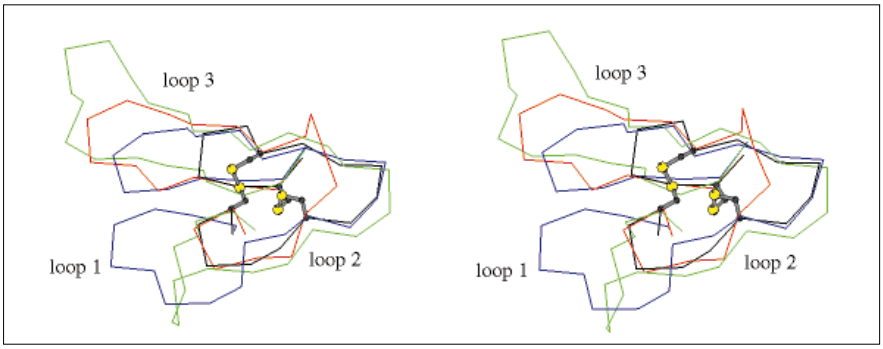
The interface between domains N and C is created by residues 12–15 and residues 40–46, respectively, and the sidechains of the amino acids are involved in the majority of the interface contacts. The sidechains of Ala14 and Tyr46 make hydrophobic interactions and hydrogen bonds are found between Ser12, Ala14, Gln15, Asp40 and Asn42. Within domain C, His28 ND1 and Asn25 ND2 interact with Tyr46 O and Val27 O, respectively. In domain N, a hydrogen bond is formed between Gln15 OE1 and Ser12 N (Fig. 2b).

Structural similarities

Hirustasin shows remarkable sequence homology with antistasin, ghilanten and guamerin [4,21]. The alignment shown in Figure 3 indicates that all cysteines are conserved and it can be anticipated that these molecules fold into similar 3D structures. The structure of hirustasin presented here, is the first 3D description of a member of this inhibitor family. As antistasin and ghilanten seem to comprise two domains, both of which have a structure similar to hirustasin, the overall structure of these inhibitors might be subdivided into four domains. However, sequence comparisons of small, cysteine-rich molecules alone must be treated with care. Decorsin and hirustasin, for example, show 26% sequence identity, and five cysteines of decorsin match perfectly with cysteines from hirustasin (data not shown); however, the sequence alignment differs significantly from the structural alignment in which only two of the three disulfide bridges are conserved (Fig. 4).

A comparison of the entire hirustasin structure with a non-redundant subset of the Brookhaven protein data base

Figure 4



Superposition of the Cα atoms of the N-terminal part of hirustasin (residues 5–23, black lines), decorsin (residues 16–39, red lines), cardiotoxin (residues 2–39, green) and cellobiohydrolase (residues 7–36, blue lines). The two disulfide-bridges crosslinking the three loops are shown as a ball-and-stick model. The superposition was calculated with the program SUPERIMPOSE [23] and generated with the program MOLSCRIPT [46].

Table 1

Equivalent cysteine residues, loop lengths and cysteine Cα distances in proteins that have the *abab* consensus motif.

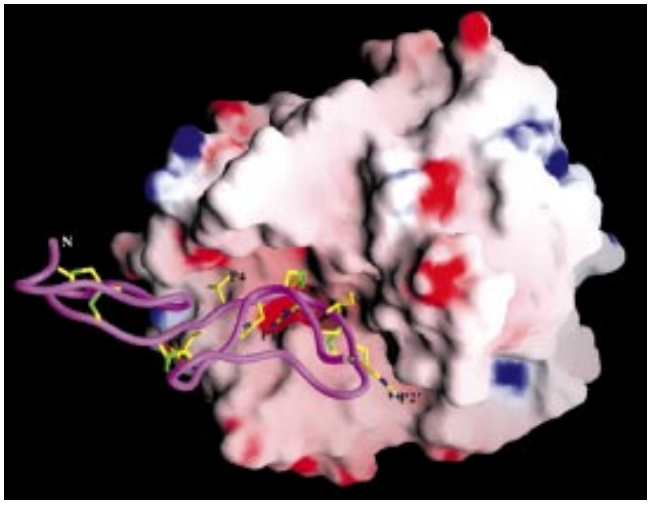
Protein		Disulfide *a*		Disulfide *b*		Length of†			
	PDB code*	Cys_1	Cys_3	Cys_2	Cys_4	Loop 1	loop 2	Loop 3	ΔCα‡ (Å)
Hirustasin		6	17	11	22	5	6	5	–
Cardiotoxin	1TGX	3	21	14	38	11	7	17	0.68
Neurotoxin-I	1NTN	3	21	14	42	11	7	21	0.69
Fasciculin 1	1FAS	3	22	17	39	14	5	17	0.38
CD59	1ERG	3	26	19	39	16	7	13	0.76
Cellobiohydrolase I	2CBH	8	25	19	35	11	6	10	0.65
Carboxypeptidase A inhibitor	4CPA	12	27	18	34	6	9	7	0.47
Cucurbita maxima trypsin inhibitor	1PPE	10	22	16	28	6	6	6	0.46
Hirudin	1HIC	16	28	22	39	6	6	11	0.50
Decorsin	1DEC	17	27	22	38	5	5	11	0.50
Average						9	6	12	0.56
(σ)						(3.8)	(1.1)	(5.0)	(0.12)

*The entry codes for the Brookhaven Protein Database. †Length of loop gives the number of residues between Cys_1–Cys_2 (loop 1), Cys_2–Cys_3 (loop 2) and Cys_3–Cys_4 (loop 3), including one cysteine. ‡Rms deviation between the specified Cα atoms from Cys_1, Cys_2, Cys_3 and Cys_4 superimposed onto Cys6, Cys11, Cys17 and Cys22 from hirustasin.

[22,23] showed no significant similarity with any of the tested structures. After subdividing the hirustasin structure into the two domains, domain C was shown to be unrelated to any of the structures in the database. Therefore, domain C represents a novel structural motif for cysteine-rich proteins. The comparison also revealed that domain N is structurally related to cardiotoxin, neurotoxin-I, fasciculin 1, the extracellular region of human complement regulatory protein CD59, the C-terminal domain of cellobiohydrolase I, and decorsin [2,24–28]. As it has already been reported that hirudin [29] is similar to decorsin [2], and that potato carboxypeptidase A inhibitor (CPI) [30] and the trypsin inhibitor from squash seeds (CMTI-I) [31] are similar to the C-terminal domain of cellobiohydrolase I [28], these structures were included in the comparison (Table 1). They vary in lengths between 29 (CMTI-I) and 77 residues (CD59), but possess the same *abab* pattern with an orientation of disulfide bridges identical to that present in the N-terminal part of hirustasin (Fig. 4). All the listed proteins contain additional residues either at the N terminus (decorsin and hirudin) or at the C terminus (hirustasin, cardiotoxin, neurotoxin-I, fasciculin 1, CD59, CPI, CMTI-I and cellobiohydrolase I) of the Cys_1-$loop_1$-Cys_2-$loop_2$-Cys_3-$loop_3$-Cys_4 motif. The N- and C-terminal extensions are often crosslinked to the consensus motif by additional disulfide bridges. The similarities between CD59 and snake venom toxins, between cellobiohydrolase I, CPI and CMTI-I, and between decorsin and hirudin have been described previously [2,27,28,31].

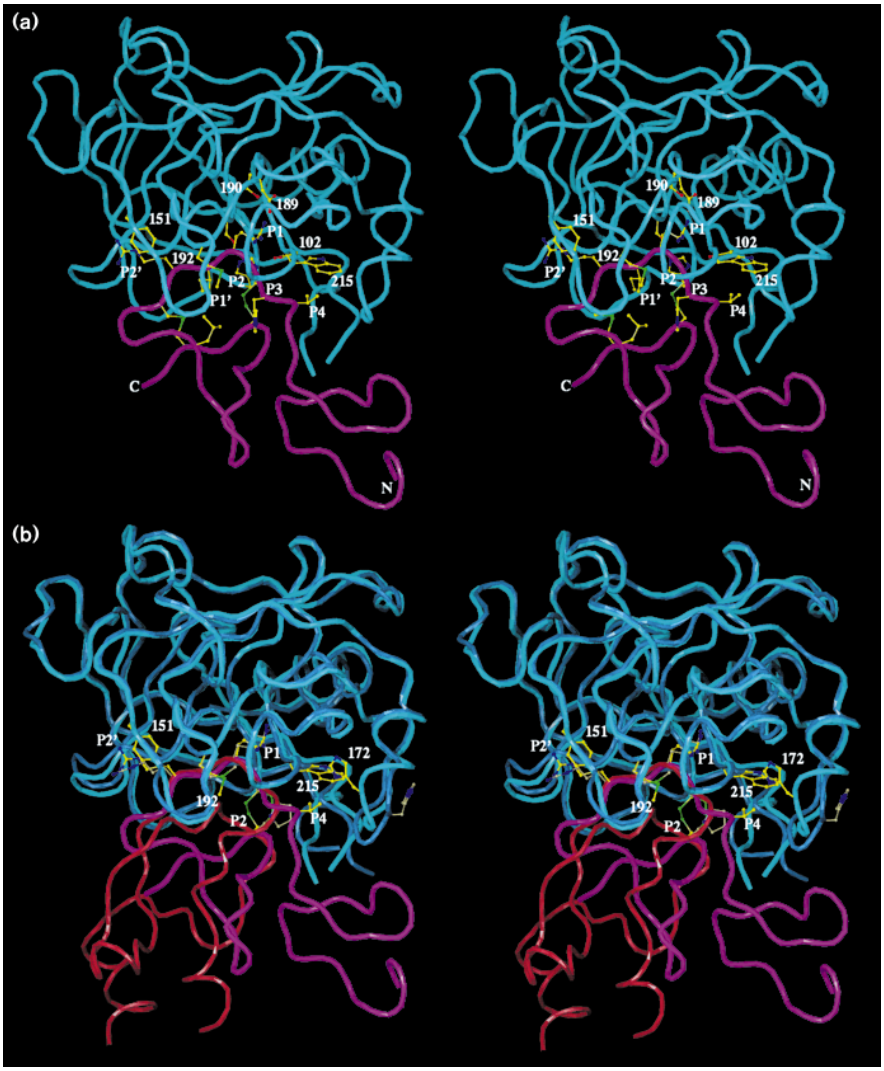
Although the lengths of the three loops between the disulfide bridges vary between 5 and 21 residues, the relative orientations of the four cysteine residues are conserved. The average Cα distance for a pairwise least squares superposition is approximately 0.6Å for the four cysteine Cα atoms. The consensus motif of the 10 structures is a left-handed helical turn in which the subsequent repeats are crosslinked by two disulfide bridges. Domain N of hirustasin is the most compact structure in any of the structures listed in Table 1. Within this group the three loops have average lengths of 9, 6 and 12 residues, respectively, with loop 1 and loop 3 being significantly more variable in length than loop 2. All proteins listed in Table 1 are inhibitors, except for the C-terminal domain of cellobiohydrolase. As domain N of hirustasin is not involved

Figure 5



Molecular surface of kallikrein colored according to the electrostatic potential. Red and blue indicate negative and positive potentials, respectively. Hirustasin is shown as a tube. The disulfide bridges and the residues in the primary binding loop are indicated as stick models. The figure was generated with the program GRASP [47].

Figure 6



The hirustasin–kallikrein complex. (a) Stereo plot of the complex between hirustasin (magenta) and kallikrein (light blue). The mainchain is depicted as a tube. Some kallikrein residues at the interface and residues P4–P2′ from hirustasin are indicated. (b) Superposition of the kallikrein(dark-blue)–aprotinin(red) [16] and kallikrein (light-blue)–hirustasin(magenta) complexes based on the kallikrein residues. Sidechains are color-coded. Residues with yellow and khaki carbon atoms belong to the kallikrein–hirustasin and kallikrein–aprotinin complexes, respectively.

in inhibiting tissue kallikrein, and because of the striking structural similarity to other cysteine-rich proteins, it is very probable that domain N has an additional function. Obviously, although these structures exhibit a similar motif, the interactions with the target enzymes are mapped to different parts of the structures.

Enzyme–inhibitor complex

Hirustasin binds in the active site pocket of kallikrein which is formed by residues 39–42, 57–58, 190–196 and 214–218. As in many other serine proteases, the active site is located in the cleft between the two kallikrein domains [20]. Interactions between the inhibitor and the protease are restricted to hirustasin domain C, particularly to the primary binding loop (residues 26–33). As hirustasin has an elongated shape and the primary binding loop is located at one of its corners, domain N of hirustasin is pointing in the direction of residues 169–172 of the enzyme (Figs 5,6).

The interface between hirustasin and kallikrein covers an area of 940 Å^2 thus comprising 26% of the surface of hirustasin. Binding of the inhibitor benefits from the opposite electrostatic potentials in the active site of the protease and on the surface of the inhibitor. Hirustasin carries a net load of three positive charges that exert a positive potential around the primary binding loop. In the complex, this area matches onto an area of negative potential in the active site of the protease (Fig. 5). This negative potential arises from Asp189 and Asp194, the only charged residues not located on the surface of the molecule. The sidechain of Asp194 is close to the active site and forms an internal salt bridge with the N terminus of the protease. Asp189 is located at the bottom of the P1 pocket: The closest distance between Asp189 and the inhibitor is 3.5 Å (Asp189 OD1–Arg30 NH2).

The scissile bond between Arg30 and Ile31 of hirustasin is sandwiched between the sidechains of Met192 and

Table 2

Hydrogen bonds in the hirustasin–kallikrein and aprotinin–kallikrein binding sites.

	Hirustasin–kallikrein		Aprotinin–kallikrein	
Atom in kallikrein	Atom in hirustasin	Distance (Å)	Atom in aprotinin	Distance (Å)
Gln A41 O	Arg I32 N	3.2	Arg 17 N	2.7
Asp B189 OD1	Arg I30 NH2	3.5	–	–
His B217 O	Arg I30 NH2	2.8	–	–
Asp B189 OD2	Arg I30 NH1	3.7	Lys 15 NZ	3.4
Thr B190 OG1	Arg I30 NH1	3.0	Lys 15 NZ	2.7
Thr B190 O	Arg I30 NH1	3.0	Lys 15 NZ	2.7
Ser B226 OG	Arg I30 NH1	3.3	Lys 15 NZ	2.9
Gly B193 N	Arg I30 O	2.9	Lys 15 O	2.8
Asp B194 N	Arg I30 O	3.4	Lys 15 O	3.2
Ser B195 N	Arg I30 O	3.1	Lys 15 O	2.9
Ser B214 O	Arg I30 N	3.0	Lys 15 N	2.9
Gly B216 N	His I28 O	3.0	Pro 13 O	2.7
Gly B216 O	His I28 N	2.8	–	–
His B217 NE2	Glu I26 OE2	3.0	–	–
Thr B218 N	Glu I26 O	3.0	–	–
Tyr B99 OH	–	–	Arg39 NE	2.8

Ser195. The distance between Ser195 OG and Arg30 C is rather short for a non-bonding interaction (2.8 Å); nevertheless, the electron density implies a planar rather than a tetrahedral configuration at Arg30 C, indicating that the scissile bond is not attacked by Ser195 OG. The carbonyl oxygen is pointing into the oxyanion hole and forms hydrogen bonds with Ser195 N and Gly193 N (Table 2).

His57, Asp102 and Ser195 comprise the catalytic triad of kallikrein. The side chain of His57 forms hydrogen bonds with the sidechains of Ser195 and Asp102 (Ser195 OG–His57 NE2, 2.7 Å; His57 ND1–Asp102 OD2, 2.9 Å). His57 and Asp102 polarize the sidechain of Ser195 for nucleophilic attack at Asp30 C. In the complex with hirustasin, these residues are shielded from solvent by the inhibitor disulfide bridge Cys29–Cys48 and by Phe94 and Leu95A from the protease.

Most of the polar interactions between the inhibitor and the protease are formed by Arg30 which fits into the P1 pocket (Figs 5,6a). All polar atoms of this residue are recognized either directly by the protease or by water molecules (Arg30 NE–Wat131–Trp141 O). As already observed in the kallikrein–aprotinin complex [16], the residues preceding the scissile bond are extensively recognized by the protease. Residues 26–30 of hirustasin and residues 214–218 of kallikrein form an antiparallel β sheet. The sidechain of Glu26 (equivalent to P5) forms a salt bridge with the sidechain of His217 on the kallikrein surface. The P1′ and P2′ residues Ile31 and Arg32 are bound between loops 39–41 and 169–198 of kallikrein. The guanidinium group of Arg32 is aligned parallel to the benzene ring of Phe153. As the guanidinium group is probably positively charged, it can interact with the π orbitals from the aromatic sidechain. Interactions between positively charged ions and aromatic groups are known to be relatively strong [32]. The secondary binding loop (residues 44–50) of hirustasin is 4–5 Å away from the active site and does not interact directly with the protease. Only the sidechain of Glu45 forms a water-mediated interaction with the protease.

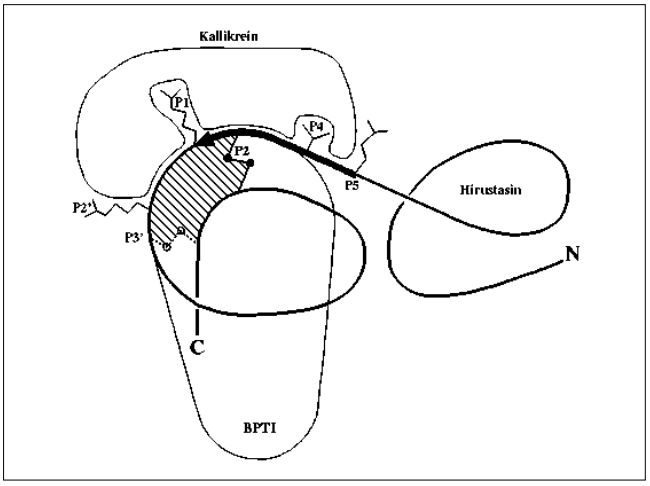
Hydrophobic interactions seem to play an important role for the recognition of hirustasin (Fig. 6a). The sidechain of Met192 from kallikrein packs against the hydrophobic ring of hirustasin formed by the disulfide bridges Cys29–Cys48 and Cys33–Cys50, and the sidechains of Leu37 and His28 (P3). Residues Trp215, His172, Pro173 and Asp174 of kallikrein create a cavity that accommodates the sidechain of Val27 (P4) of hirustasin. The size and the hydrophobicity of this cavity favor the binding of a small hydrophobic sidechain.

As already observed in the comparison between the uncomplexed kallikrein structure with the kallikrein–aprotinin complex [15,16], the bulky disulfide bridge in the P2 position of hirustasin forces the sidechain of Tyr99 of kallikrein to move away from its position. This movement is accompanied by a general reorganization of the preceding loop region (residues 95–99), which is also involved in a crystal contact.

Comparison with related enzyme–inhibitor complexes

The crystal structure of the kallikrein–aprotinin complex has been solved at 2.5 Å resolution [16]. The structure of kallikrein in complex with aprotinin is very similar to the structure of kallikrein in complex with hirustasin (rms deviation Cα = 0.58 Å). In the hirustasin–kallikrein complex, the most significant difference is observed for residues 169–174 which contribute to the pocket that recognizes the P4 valine. Asp174 has moved 1 Å away from the inhibitor to

Figure 7



Schematic superposition of hirustasin (bold line) with aprotinin (gray surface) similar to Figure 6b. Two disulfide bridges from the C-terminal part that are conserved and absent in the two molecules are sketched in solid and broken lines, respectively. The hydrophobic region is striped. Residues in P3′, P2′, P1, P2, P4 and P5 are labeled. The P1 and P4 pockets in kallikrein are indicated.

provide sufficient space for the P4 valine. Residues 171–173 move in the opposite direction and close off the P4 pocket from the solvent. The largest shift is seen for His172, the Cα of which moves 4 Å towards the inhibitor. In the structure of the uncomplexed kallikrein [15], this loop adopts a conformation similar to that in the complex with hirustasin. This difference indicates the ability of loop 169–174 to adapt its conformation to the substrate or inhibitor. In the aprotinin complex, the P4 pocket is not occupied because the aprotinin mainchain makes a sharp turn after Pro13, away from the enzyme (Fig. 6b). In aprotinin, the primary binding loop contains recognition sites P2′–P3, but in hirustasin the antiparallel β-sheet is extended at the N-terminal end by two residues (P4, Val27; P5, Glu26).

In hirustasin, the P1 residue is an arginine, but in aprotinin it is a lysine. Although the arginine in hirustasin is longer than the lysine in aprotinin, it cannot get closer to Asp189 at the bottom of the binding pocket, most probably because of the size of the guanidinium-group and its ability to form additional hydrogen bonds. Arg30 NH1 from hirustasin matches perfectly onto Lys15 NZ from aprotinin and forms a slightly longer hydrogen bond with the protease. Arg30 NH2 forms additional hydrogen bonds with His217 O and Asp189 OD1 (Table 2). Most of the sidechain conformations in kallikrein are similar in the two complexes. Striking differences are observed for the side chain of Met192. In the complex with hirustasin, Met192 CE is lying in a hydrophobic pocket formed by the inhibitor. This movement is supported by the different P1 sidechain. In the complex with aprotinin, Met192 CE points towards Lys15 CE. Due to the bulky guanidinium-group in hirustasin, Met192 CE can no longer be accommodated in this position.

Figure 7 shows a schematic superposition of hirustasin and aprotinin. As in many other protease inhibitors, the overall structures of the inhibitors are totally unrelated, but the mainchain conformations of the primary binding loops are identical (reviewed in [20]). In hirustasin, residues 28–32, which are equivalent to the P3–P2′ sites, match perfectly onto residues 13–17 from aprotinin. Residues in P3–P1 adopt a β-sheet conformation and form similar hydrogen bonds with the protease (Table 2). Among these residues, only arginine at P2′ and cysteine at P2 are identical. The interactions between the guanidinium group from the arginine in P2′ (Arg32 in hirustasin, Arg17 in aprotinin) and the benzene ring of Phe151 from kallikrein are conserved in both structures. The cysteine residue in P2 is involved in a disulfide bridge with the secondary binding loop (Cys29–Cys48 in hirustasin, Cys14–Cys38 in aprotinin).

Interestingly, and in contrast to antistasin and ghilanten, hirustasin does not inhibit factor Xa. A docking study based on the structure of the uncomplexed human factor Xa [33] superimposed onto the kallikrein structure revealed that most of the interactions responsible for correct binding could be formed by factor Xa as well. Hirustasin residues that form sidechain-specific interactions with kallikrein are P5, P4, P1 and P2′. From the alignment shown in Figure 3, it is evident that in antistasin and ghilanten P1 is always an arginine and P4 is either a valine or an isoleucine. Except for domain 1 of antistasin, there is always a negatively charged residue in P5. The greatest sequence variability in the primary binding loops is observed for P2′. Indeed, only hirustasin contains an arginine that is able to interact with the benzene sidechain of Phe151. As factor Xa contains no aromatic residue at this position, the interaction between Arg32 and Phe151 is specific for the recognition of tissue kallikrein by hirustasin. A similar interaction is also found in the complex of aprotinin with kallikrein [16] and with trypsin [34]. The relevance of an arginine at P2′ is also supported by the observation that wild-type aprotinin is a very weak inhibitor for factor Xa [35].

Biological implications

Hirustasin acts as a potent inhibitor of tissue kallikrein and has a potential medical application in those diseases in which the tissue kallikrein–kinin system or a tissue kallikrein-related protease plays a major role. Kallidin or lysyl-bradykinin belong to the kinin class of biologically or pharmacologically highly potent tissue hormones, and are released from their natural target substrate, the kininogens, specifically by tissue kallikreins. Kinins bind to their receptors on specific target cells and exert a broad variety of biological activities, such as vasodilatation,

blood pressure reduction, smooth muscle relaxation or contraction, pain and inflammation. Kinin generation is primarily determined by the activity and availability of kallikrein, because the level of kininogen is not the rate-limiting factor. Kallikrein levels are controlled by the rate of proenzyme synthesis and activation, as well as by inhibition and elimination of the kallikrein–inhibitor complexes. The transcription of tissue kallikreins is subjected to regulation by a number of hormones and transcription factors. At the post-translational level, the activity and metabolism of kallikreins are modulated by endogenous kallikrein-binding proteins. The human glandular kallikrein-1 is very similar to the prostate-specific antigen (PSA), which is the most sensitive marker available for monitoring prostate cancer progression and response to therapy. PSA, which does not liberate kinins from kininogens, seems to have biological activities involved in tumor growth and metastasis [36]. The human tissue kallikrein-1 probably has an additional physiological role in the prostate, in seminal fluid or at metastatic sites. Elevated levels of tissue kallikrein have been found in a human colon carcinoma cell line and in human breast cancer cells [37, 38].

The structure of the hirustasin–kallikrein complex presented here reveals that this inhibitor differs from the other known members of various inhibitor families in its three-dimensional (3D) structure and disulfide-bridging pattern. On the basis of sequence identity and the conservation of the disulfide-bridging pattern, it is likely that a new class of leech-derived protease inhibitors, including antistasin, guamerin and ghilanten, have a 3D structure similar to hirustasin. As these molecules all have medical applications, the structural information can be used for the rational design of low molecular weight inhibitors.

Hirustasin has a compact structure that is subdivided into two domains. Although both domains show a similar pattern of cysteine residues, the exact orientation of disulfide bridges is different. As hirustasin lacks a hydrophobic core, the disulfide bridges maintain a defined 3D structure and serve as a scaffold for the primary binding loop that interacts with the protease. The primary binding loop adopts a canonical conformation that is found in many other protease inhibitors. The protease recognizes this loop as an antiparallel **β** sheet, using sidechain-specific interactions that involve residues P5, P4, P1 and P2′ of the inhibitor. Although the P1 arginine contributes most of the hydrogen bonds between the protease and the inhibitor, the P2′ arginine seems to be responsible for the discrimination between tissue kallikrein and factor Xa.

The N-terminal domain of hirustasin does not interact with the protease but it is structurally similar to a number of other proteins such as cardiotoxin, neurotoxin-I, fasciculin 1, the extracellular region of human complement regulatory protein CD59, the C-terminal domain of cellobiohydrolase I, carboxypeptidase A inhibitor, trypsin inhibitor from squash seeds, hirudin and decorsin. The consensus motif can be described as a left-handed helical turn that is connected by at least two disulfide bridges. The lengths of the three loops connecting the cysteine residues differ between 5 and 21 residues.

Materials and methods

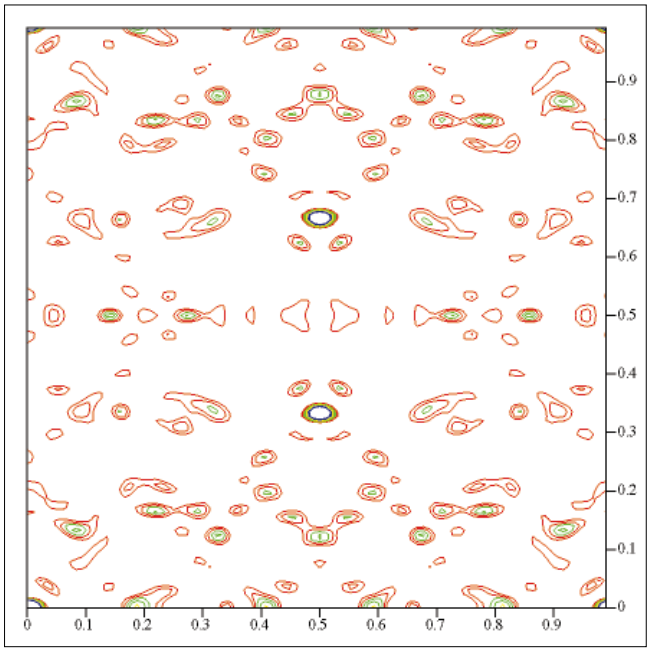
Crystallization

Pure recombinant hirustasin was purified as described by Di Marco *et al.* [21]. One aliquot (50 μl) of hirustasin was mixed with a 50 μl aliquot of kallikrein to final kallikrein and hirustasin concentrations of 0.50 mM and 0.85 mM, respectively, in 20 mM Tris-HCl, pH 8.0. Crystals were grown using the 'hanging drop' method. The crystallizing solution consisted of 23% polyethyleneglycol (PEG) 2000 monomethyl ether, 0.18 M ammonium sulfate and 3.5% dioxane in 0.1 M sodium acetate buffer, pH 4.6. Crystals to a maximum size of $1\times0.5\times0.1$ mm^3 were obtained after 2–3 days at room temperature. These crystals were stable only for one week, after which they started to change in morphology and birefringence. The 1:1 ratio of hirustasin and kallikrein in the crystals was confirmed by SDS-PAGE and by RP-HPLC (data not shown).

Data collection

For data collection, crystals were transferred from the crystallization drop to a stabilizing solution consisting of 34% polyethylenglycol (PEG) 2000 monomethyl ether and 0.2 M ammonium sulfate, in 0.1 M sodium acetate buffer, pH 4.6. Two native data sets were collected at room temperature on similar image-plate systems (Mar Research), one mounted at the

Figure 8



Harker section (w = 1/2) of the native Patterson map, calculated for the resolution range 20.0–2.4 Å (u, horizontal; v, vertical). The symmetry-related peaks at (1/2, 1/3, 1/2) and (1/2, 2/3, 1/2) possess 50.5% of the origin peak height.

Table 3

Statistics of data sets.

	Data set 1	Data set 2	Combined data sets	
			All data	Last shell
Unit cell parameters (Å)	a = 69.19, b = 86.02,	a = 69.57, b = 85.97,	–	–
	c = 116.63	c = 117.23	–	–
Wavelength (Å)	1.54	0.875	–	–
Detector distance (mm)	120	200	–	–
Frame / exposure (sec^{-1})	0.5° / 360	1.0° / 180	–	–
Resolution (Å)	24.5–2.4	24.5–2.4	24.5–2.4	2.45–2.40
R_{sym} (%)*	12.5	12.8	12.6	47.4
Observations	52 768	42 129	96 170	4 989
Unique reflections	–	–	27 511	1 623
Completeness (%)	–	–	97.9	98.6
Percentage $[I] \geq 2\sigma$ (%)	–	–	56.8	32.9
Multiplicity	–	–	3.5	3.1

*$R_{sym} = \sum_{hkl} \sum_{i} |I_{hkl,i} - < I_{hkl} > | / \sum_{hkl} \sum_{i} |I_{hkl,i}|$.

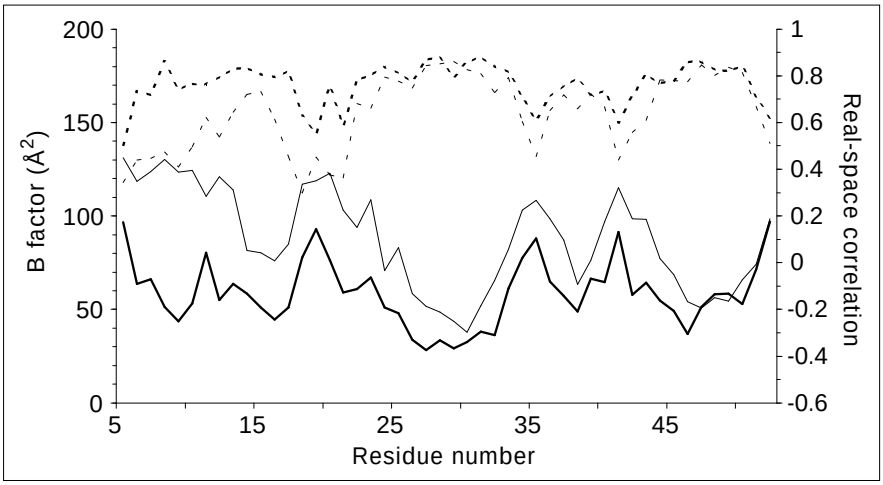
Swiss/Norwegian beam-line at the European Synchrotron Radiation Facility (ESRF, Grenoble) and the other on an in-house FR591 rotating anode system (Nonius). Statistics of the individual and the combined data sets are given in Table 3. Both data sets were processed with MARXDS [39] and combined using MARSCALE. Only the first 50° of data were useful because the crystals suffered strong radiation damage. Because of the radiation damage and high mosaicity, the data were of poor quality as indicated by a R_{sym} of 12.6%. The average intensity of the diffraction pattern was weaker in the direction of the c* axis than in the a*/b* plane. Because of the absence of the odd reflections along h, k and l, we assigned the orthorhombic space-group $P2_12_12_1$ with unit-cell dimensions of a = 69.4 Å, b = 86.0 Å, c = 116.9 Å, a = b = γ = 90°. Visual inspection of the combined data set indicated that all reflections with k = 3n: h + l ≠ 2n were either weak or absent.

Structure solution and refinement

The structure was solved by molecular replacement (program AMoRe) [40] using the 2.05 Å crystal structure of porcine kallikrein (Brookhaven accession code: 2PKA) as a search model. We assumed two hirustasin–kallikrein complexes per asymmetric unit with a packing parameter of 2.77 $Å^3\ Da^{-1}$. However, the cross-rotation functions calculated for various resolution ranges, showed only one significant peak. Using data between 8.0 and 3.0 Å resolution and a Patterson sphere of 20 Å, an 11σ peak was found for the rotation α = 112.21°, β = 31.25° and γ = 167.40°. The absence of a second solution was in agreement with the absence of any significant peaks in the self-rotation function and with the results of a native Patterson map which showed one significant peak at (1/2, 1/3, 1/2) (Fig. 8).

Preliminary translation functions calculated in space-group $P2_12_12_1$ yielded promising correlation and R_f values, but the packing of the second molecule was always very poor. As the translation vectors observed in the native Patterson map can generate pseudo-extinction patterns along a* and c*, we examined the alternative space-group $P2_12_12$. The translation function for the first molecule gave a significant solution for a = 0.2510, b = 0.3205 and c = 0.4631 (r = 48.9%, R_f = 54.4%, 10.0–3.0 Å resolution). For the second molecule, we found two conspicuous solutions of nearly equivalent height (1. solution: a = 0.7514, b = 0.6526, c = 0.9635, r = 67.2%, R_f = 47.5%; 2. solution: a = 0.7513, b = 0.9882, c = 0.9631, r = 66.9%, R_f = 48.3%, 10.0–3.0 Å resolution). These two solutions differed only in Δb = 0.33. Because the F_o–F_c map showed clearly the bound inhibitor, we took this as an indication for the correctness of solution 1.

Figure 9



Average B factors (continuous lines) and real-space correlation values (broken lines) for the two hirustasin chains (thick lines, residues I5–I52; thin lines, residues J5–J52) are shown plotted over the residue number.

The initial R-factor for the correctly positioned search-model was 50.2% for all data between 7.0 and 2.4 Å resolution.

The twofold averaged σ_A-weighted electron-density maps (program AVE [41] and SIGMAA [42]) were sufficiently clear to trace the hirustasin main chain and to assign the disulfide bridges. We subsequently refined the model to a working R-factor of 31.2% (free R factor = 42.2%) using X-PLOR [43] and O [44]. Because of the anisotropy of the data set, we applied overall anisotropic B factor refinement [45]. Individual inspection of the two complexes showed good electron density for the first molecule, but very poor density for the second. This led us to check whether we could solve the structure by molecular replacement in space-group $P2_12_12$ with a = 116.9 Å, b = 86.0 Å, c = 69.4 Å, $\alpha = \beta = \gamma = 90°$ using the pre-refined complex. The cross-rotation function yielded a 13.8σ peak at $\alpha = 66.00°$, $\beta = 154.67°$ and $\gamma = 93.50°$ (8.0–3.5 Å resolution and a Patterson sphere of 20 Å). The translation function clearly showed solutions for the first complex at a = 0.2506, b = 0.2623 and c = 0.000 (r = 43.4%, R_f = 57.2%, 8.0–3.5 Å resolution) and for the second complex at a = 0.7507, b = 0.5964 and c = 0.5020 (r = 86.6%, R_f = 30.3%, 8.0–3.5 Å resolution). The R factor for the correctly positioned search-model was 34.2% (8.0–2.4 Å resolution). It was subsequently reduced to 20.5% (free R factor = 31.1%) for all data between 8.0 and 2.4 Å. The rms deviations for bond lengths and angles are 0.011 Å and 1.72°, respectively. During refinement, the second complex was transferred into the asymmetric unit of the first complex by applying the symmetry operator (-x, -y, z) and the translation (1,1,0). The final structure of the hirustasin–kallikrein complex contains two copies of residues 16–247 from the protease and residues 5–52 from the inhibitor. All mainchain dihedral angles fall into allowed regions of the Ramachandran diagram. The average B factor for all 4584 atoms is 46.2 Å^2. Among them there are 304 water molecules with an average B factor of 53.6 Å^2. The average B factors of the two kallikrein molecules are 37.9 Å^2 (41.8 Å^2) and of the two hirustasin molecules 61.7 Å^2 (88.1 Å^2), respectively. The average B factors and real-space correlation values [44] for hirustasin are plotted over the residue number in Figure 9.

Database searches

To check if the structure of hirustasin is similar to any known structure, it was superimposed on all structures from a non-redundant subset of the Brookhaven Protein Database. This database contained 721 structures, all of which have less than 35% sequence identity [22]. For the automatic calculation of optimal superpositions the program SUPERIMPOSE [23] was used. Residues 5–23 and 23–52 of hirustasin were checked individually. Significant structural similarity was indicated only for the N-terminal part by best-fit values between 3.9σ and 2.8σ.

Accession numbers

The coordinates of the hirustasin–kallikrein complex have been deposited at the Brookhaven Protein Databank (ID code 1H1A).

Acknowledgements

We are grateful to Bayer AG (Drs. Hörlein, Nöthen and Stadler) for providing us with highly purified porcine tissue kallikrein (1599 KU mg^{-1} protein).

References

1. Walsmann, P. & Markwardt, F. (1985). On the isolation of the thrombin inhibitor hirudin. *Thromb. Res.* 40, 563–569.
2. Kretzel, A.M., Wagner, G., Seymour-Ulmer, J. & Lazarus, R.A. (1994). Structure of the RGB protein decorsin: conserved motif and distinct function in leech proteins that affect blood clotting. *Science* 264, 1944–1947.
3. Mazur, P., Henzel, W.J., Seymour, J.L. & Lazarus, R.A. (1991). Ornatins: potent glycoprotein IIB-IIIa antagonists and platelet aggregation inhibitors from the leech *Placobdella ornata. Eur. J. Biochem.* 202, 1073–1082.
4. Söllner, C., Mentele, R., Eckerskorn, C., Fritz, H. & Sommerhoff, C.P. (1994). Isolation and characterization of hirustasin, an antistasin-type serine protease inhibitor from the medical leech *Hirudo medicinalis. Eur. J. Biochem*. 219, 937–943.
5. Tuszynski, G.P., Gasic, T.B. & Gasic, G.J. (1987). Isolation and characterization of antistasin, an inhibitor of metastasis and coagulation. *J. Biol. Chem.* 262, 9718–9723.
6. Nutt, E., *et al*., & Simpson, E. (1988). The amino acid sequence of antistasin: a potent inhibitor of factor Xa reveals a repeated internal structure. *J. Biol. Chem.* 263, 10162–10167.
7. Condra, C., Nutt, E., Petroski, J.C., Simpson, E., Friedman, P.A. & Jacobs, J.W. (1989). Isolation and structural characterization of a potent inhibitor of coagulation factor Xa from the leech *Haementeria ghilianii. Thromb. Haemostas.* 61, 437–441.
8. Jung, H.I., Kim, S.I., Ha, K.-S., Joe, C.O. & Kang, K.W. (1995). Isolation and characterization of guamerin, a new human leukocyte elastase inhibitor from *Hirudo nipponia. J. Biol. Chem.* 270, 13879–13884.
9. Seemüller, U., Dodt, J., Fink, E. & Fritz, H. (1986). Protease inhibitors of the leech *Hirudo medicinalis* (hirudins, bdellins, eglins). In *Protease Inhibitors*. (Barrett, A.J. & Salvesen, G., eds), pp. 337–359, Elsevier, Amsterdam.
10. Bhoola, K.D., Figueroa, C.D. & Worthy, K. (1992). Bioregulation of kinins: kallikreins, kininogens, and kininases. *Pharmacol. Rev.* 44, 1–80.
11. Rawlings, N.D. & Barrett, A.J. (1994). Families of serine peptidases. *Meth. Enzymol.* 244, 19–61.
12. Müller-Esterl, W., Iwanga, S. & Nakaniski, S. (1986). Kininogens revisited. *Trends Biochem. Sci.* 11, 336–339.
13. Chao, J. & Chao, L. (1996). Functional analysis of human tissue kallikrein in transgenic mouse models. *Hypertension* 27, 491–494.
14. Fiedler, F. & Leysath, G. (1979). Kinins II. In *Advances in Experimental Medicine and Biology 120A*. (Fujii, S., Moriya, H. & Suzuki, T., eds), pp. 261–271, Plenum Press, New York & London.
15. Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G. & Bartunik, H. (1983). Refined 2 Å X-ray crystal structure of porcine pancreatic kallikrein A, a specific trypsin-like serine protease. *J. Mol. Biol.* 164, 237–258.
16. Chen, Z. & Bode, W. (1983). Refined 2.5 Å X-ray structure of the complex formed by porcine kallikrein A and the bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 164, 283–311.
17. Schulz, G.E. & Schirmer, R.H. (1985). Structural domains. In *Principles of Protein Structure*. (Schulz, G.E. & Schirmer, R.H. eds), pp. 84–95, Springer-Verlag, New York, Berlin Heidelberg & Tokyo.
18. Benham, C.J. & Jafri, S. (1993). Disulfide bonding patterns and protein topologies. *Protein Sci.* 2, 41–54.
19. Huang, K., Strynadka, N.C.J., Bernard, V.D., Peanasky, R.J. & James, M.N.G. (1994). The molecular structure of the complex of *Ascaris* chymotrypsin/elastase inhibitor with porcine elastase. *Structure* 2, 679–689.
20. Bode, W. & Huber, R. (1992). Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem*. 204, 433–451.
21. Di, Marco, S., *et al*., & Grütter, M.G. (1997). Recombinant hirustasin: production in yeast, crystallization and interaction with serine proteases. *Protein Sci.* 6, 103–118.
22. Hobohm, U. & Sander, C. (1994). Enlarged representative set of protein structures. *Protein Sci.* 3, 522–524.
23. Diederichs, K. (1995). Structure superposition of proteins with unknown alignment and detection of topological similarity using a six-dimensional search algorithm. *Proteins* 23, 187–195.
24. Bilwes, A., Rees, B., Moras, D., Ménez, R. & Ménez, A. (1994). X-ray structure of 1.55 Å of toxin g, a Cardiotoxin from *Naja nigricollis* venom. *J. Mol. Biol.* 239, 122–136.
25. Nickitenko, A.V., Michailov, A.M., Betzel, C. & Wilson, K.S. (1993). Three-dimensional structure of neurotoxin-1 from *Naja naja oxiana* venom at 1.9 Å resolution. *FEBS Lett.* 230, 111–117.
26. Du le, M.H., Marchot, P., Bougis, P.E. & Fontecilla-Camps, J.C. (1992). 1.9 Å resolution structure of fasciculin 1, an anti-acetylcholinesterase toxin from green mamba snake venom. *J. Biol. Chem.* 267, 22122–22130.
27. Fletcher, C.M., Harrison, R.A., Lachmann, P.J. & Neuhaus, D. (1994). Structure of a soluble, glycosylated form of the human complement regulatory protein CD59. *Structure* 2, 185–199.
28. Kraulis, P.J., *et al*., & Gronenborn, A.M. (1989). Determination of the three-dimensional solution structure of the C-terminal domain of Cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry* 28, 7241–7257.
29. Szyperski, T., Guntert, P., Stone, S.R. & Wüthrich, K. (1992). The NMR solution structure of hirudin (1–51) and comparison with corresponding three-dimensional structures determined using the complete 65-residue hirudin polypeptide chain. *J. Mol. Biol.* 228, 1193–1205.

30. Rees, D.C. & Lipscomb, W.N. (1982). Refined crystal structure of the potato inhibitor complex of carboxypeptidase at 2.5 Å resolution. *J. Mol. Biol.* 160, 475–481.
31. Bode, W., Greyling, H.J., Huber, R., Otlewski, J. & Wilusz, T. (1989). The refined 2.0 Å X-ray crystal structure of complex formed between bovine beta-trypsin and CMTI-I, a trypsin inhibitor from squash seeds (*Cucurbita maxima*): topological similarity of carboxypeptidase A inhibitor from potatoes. *FEBS Lett.* 242, 285–292.
32. Mitchell, J.B.O., *et al*., & Singh, J. (1993). Amino/aromatic interactions. *Nature* 366, 413.
33. Padmanabhan, K., *et al*., & Kisiel, W. (1993). Structure of human des(1–45) factor Xa at 2.2 Å resolution. *J. Mol. Biol.* 232, 947–966.
34. Marquart, M., Walter, J., Deisenhofer, J., Bode, W. & Huber, R. (1983). The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Cryst. B* 39, 480–491.
35. Stassen, J.M., *et al*., & Vermylen, J. (1995). Characterisation of a novel series of BPTI-derived anticoagulants. *Thromb. Haemostas.* 74, 646–654.
36. Peehl, D.M. (1995). Prostate specific antigen role and function. *Cancer* 75, 2021–2026.
37. Chen, L.-M., Richards, G.P., Chao, L. & Chao, J. (1995). Molecular cloning, purification and *in situ* localization of human colon kallikrein. *Biochem. J.* 307, 481–486.
38. Hermann, A., Buchinger, P. & Rehbock, J. (1995). Visualization of tissue kallikrein in human breast carcinoma by two-dimensional western blotting and immunohistochemistry. *Biol. Chem. Hoppe-Seyler* 376, 365–370.
39. Kabsch, W. (1988). Evaluation of single crystal X-ray diffraction from a position-sensitive detector. *J. Appl. Cryst.* 21, 916–924.
40. Nazava, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. D* 50, 157–163.
41. Kleywegt, G.J. & Jones, T.A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Cryst. D* 50, 178–185.
42. Read, R.J. (1986). Improved fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst. A* 42, 140–149.
43. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* 235, 458–460.
44. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and location of errors in these models. *Acta Cryst. A* 47, 110–119.
45. Sheriff, S. & Hendrickson, W.A. (1987). Description of overall anisotropy in diffraction from macromolecular crystals. *Acta Cryst. A* 43, 118–121.
46. Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946–950.
47. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 281–296.