Structural Basis for Dual-Inhibition Mechanism of a Non-Classical Kazal-Type Serine Protease Inhibitor from Horseshoe Crab in Complex with Subtilisin

Rajesh T. Shenoy¹, Saravanan Thangamani¹,⁵, Adrian Velazquez-Campoy³⁴, Bow Ho², Jeak Ling Ding¹*, J. Sivaraman¹*

¹ Department of Biological Sciences, National University of Singapore, Singapore, Singapore, ² Department of Microbiology, National University of Singapore, Singapore, Singapore, ³ Institute of Biocomputation and Physics of Complex Systems (BIFI), Universidad de Zaragoza, Zaragoza, Spain, ⁴ Fundacion ARAID, Diputacion General de Aragon, Zaragoza, Spain, ⁵ Department of Pathology, Center for Biodefense and Emerging Infectious Diseases, University of Texas Medical Branch, Galveston, Texas, United States of America

Abstract

Serine proteases play an important role in host-pathogen interactions. In the innate immune system of invertebrates, multidomain protease inhibitors are important for the regulation of host-pathogen interactions and antimicrobial activities. Serine protease inhibitors, 9.3-kDa CrSPI isoforms 1 and 2, have been identified from the hepatopancreas of the horseshoe crab, Carcinoscorpius rotundicauda. The CrSPIs were biochemically active, especially CrSPI-1, which potently inhibited subtilisin (K<sub>i</sub> = 1.43 nM). CrSPI has been grouped with the non-classical Kazal-type inhibitors due to its unusual cysteine distribution. Here we report the crystal structure of CrSPI-1 in complex with subtilisin at 2.6 Å resolution and the results of biophysical interaction studies. The CrSPI-1 molecule has two domains arranged in an extended conformation. These two domains act as heads that independently interact with two separate subtilisin molecules, resulting in the inhibition of subtilisin activity at a ratio of 1:2 (inhibitor to protease). Each subtilisin molecule interacts with the reactive site loop from each domain of CrSPI-1 through a standard canonical binding mode and forms a single ternary complex. In addition, we propose the substrate preferences of each domain of CrSPI-1. Domain 2 is specific towards the bacterial protease subtilisin, while domain 1 is likely to interact with the host protease, Furin. Elucidation of the structure of the CrSPI-1: subtilisin (1:2) ternary complex increases our understanding of host-pathogen interactions in the innate immune system at the molecular level and provides new strategies for immunomodulation.

Citation: Shenoy RT, Thangamani S, Velazquez-Campoy A, Ho B, Ding JL, et al. (2011) Structural Basis for Dual-Inhibition Mechanism of a Non-Classical Kazal-Type Serine Protease Inhibitor from Horseshoe Crab in Complex with Subtilisin. PLoS ONE 6(4): e18838. doi:10.1371/journal.pone.0018838

Editor: Petri Kursula, University of Oulu, Germany

Received September 21, 2010; Accepted March 21, 2011; Published April 26, 2011

Copyright: © 2011 Shenoy et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Academic Research Fund (ARF), National University of Singapore (NUS), grant no. R154000438112. A.V.-C was supported by the Spanish Ministry of Science and Innovation (grant number BFU2010-19451); the Diputacion General de Aragon (grant number P044/09); and the Fundacion ARAID, Diputacion General de Aragon (to A.V.-C). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Rajesh T. Shenoy is a graduate scholar in receipt of a research scholarship from the NUS.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: dbsjayar@nus.edu.sg (JS); dbsdjl@nus.edu.sg (JLD)

Introduction

Serine proteases play an important immunomodulatory role in host-pathogen interactions. Invertebrates lack an adaptive immune system that recognizes and remembers specific pathogens [1]. As an evolutionarily conserved and ancient defense strategy, the innate immune system responds instantaneously to invading pathogens in a non-specific manner. The innate immune system in the horseshoe crab, Carcinoscorpius rotundicauda, comprises serine protease cascades that are similar to the blood coagulation, melanization and complement systems [2]. Horseshoe crab hemocytes contain granules filled with several serine protease zymogens. Upon mechanical injury or pathogen invasion, the granules are released into the extracellular milieu by exocytosis. Furthermore, clotting enzymes, in their precursor forms, are activated by a serine protease cascade that is triggered by bacterial endotoxins [3,4]. Several serine protease zymogens, including proclotting enzymes, Factor B and Factor C, are associated with the hemolymph coagulation system. The subsequent formation of the coagulation plug prevents further entry of pathogens [5,6].

In parallel, a large number of protease inhibitors are also expressed in hemocytes with varying specificities, which target the proteases of different microorganisms [7], including the serine proteases like elastase, subtilisin and protease K. These inhibitors belong mainly to the Kazal, Kunitz and 2-macroglobulin families [2] and they use a lock-and-key mechanism with their reactive site loops to mechanistically bind the active sites of the target proteases [8]. It has been speculated that these inhibitors might participate in the regulation of the hemolymph coagulation cascade [9,10]. During our recent work on subtractive ESTs (expressed sequence tags) from the Carcinoscorpius rotundicauda that was infected with Pseudomonas aeruginosa, two isoforms of the non-classical Kazal-type inhibitor, CrSPI, each of 9.3 kDa, were discovered [9]. The CrSPI-1 and CrSPI-2 (GenBank Accession numbers DQ090491 and DQ090492, respectively) isoforms share 97% sequence identity. Both isoforms are biochemically active. In our earlier study [9], we have tested representative microbial serine proteases
such as proteinase K and subtilisin, in comparison with eukaryotic serine protease, trypsin. While trypsin was inhibited only by CrSPI-2, subtilisin was found to be most susceptible for inhibition by CrSPI-1, with Ki of $1.4 \times 10^{-8}$ M. Hence subtilisin was used to maximize the exploration of the activity of CrSPI-1. Moreover by CrSPI-1, with Ki of 1.4

6

10

9 M. Hence subtilisin was used to

11

12

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45
Table 1. Crystallographic data and refinement statistics.

| Data collection* |            |            |            |            |            |
|------------------|------------|------------|------------|------------|------------|
| Unit cell parameters (Å) | a = 73.8, b = 65.1, c = 111.9 | α = 90, β = 95.44, γ = 90 |
| Space group | P2₁ |
| Resolution range (Å) | 50-2.6 (2.64-2.60) |
| Wavelength (Å) | 0.9600 |
| Observed hkl | 205529 |
| Unique hkl | 32884 |
| Completeness (%) | 99.8 (99.3) |
| Resolution range (Å) | 15-2.6 |
| Refinement and quality of model |            |            |            |            |
| Rwork (no. reflections) | 0.21 (24818) |
| Rfree (no. reflections) | 0.24 (1533) |
| Root mean square deviation | 6.3 (5.3) |
| Bond length (Å) | 0.009 |
| Bond angle (°) | 1.52 |
| Disallowed regions (%) | 0.0 |
| Favored region (%) | 93.9 |
| Allowed regions (%) | 5.0 |
| Generously allowed region (%) | 1.1 |
| Disallowed regions (%) | 0.0 |
| Average B-factors (Å²) |            |            |            |            |
| CrSPI-1 (no. of atoms) | 34.2 (5760) |
| Subtilisin (no. of atoms) | 32.4 (5760) |
| Overall protein atoms (no. of atoms) | 37.4 (6274) |
| Waters (no. of atoms) | 35.5 (168) |

The structure of recombinant CrSPI-1 in complex with subtilisin was solved by the molecular replacement method from a synchrotron dataset. The model was refined to a final R-factor of 0.21 (Rfree = 0.24) at 2.6 Å resolution with good stereochemical parameters (Table 1). The final refined model consisted of residues from Cys1 to Val73 of CrSPI-1 and Ala1 to Gln274 of subtilisin. Eight residues (Val74-Glu81) at the C-terminus of CrSPI-1 lacked interpretable electron density and were not modeled. There were three subtilisin molecules and one CrSPI-1 molecule in the asymmetric unit. Each CrSPI-1 molecule interacted with two subtilisin molecules, i.e., domain 1 and domain 2 of CrSPI-1 interacted with two independent subtilisin molecules (Fig. 1). The third subtilisin molecule of the asymmetric unit was not in a complex with CrSPI-1. This indicates that free subtilisin molecules were present in the crystallization drop due to the lack of adequate inhibitor molecules.

The conformation of the subtilisin molecules in the complex was similar to that seen in the subtilisin Carlsberg (EC 3.4.21.62) structure and structures of subtilisin complexed with small ligands [34–36]. The structure consists of a central seven-stranded parallel β-sheet with two α-helices on one side and a group of four α-helices on the other side of the central β-sheet. The catalytic triad, which consists of Ser220, His64 and Asp32, was located in the substrate-binding cleft. The three subtilisin molecules of the asymmetric unit were identical with an rmsd of 0.1 Å in a pairwise superposition of 274 CrSPI-1 atoms. This indicates that there was no significant conformational change of subtilisin molecules upon complex formation with the CrSPI-1 molecule.

Structure of CrSPI-1

The CrSPI-1 molecule comprises two domains: domain 1 from Cys1 to Glu40 and domain 2 from Leu41 to Leu83. Both domains adopted similar secondary structures (domain 1 [β1][β2][α1][β3] and domain 2 [β4][β5][α2]). The presence of a central α-helix, z1 (Glu18–Ala24) in domain 1 and z2 (Arg63-Ser68) in domain 2, and an antiparallel β-sheet in each domain are characteristic of the classical
Kazal-type inhibitors. However, CrSPI-1 also showed features of a non-classical Kazal-type serine protease inhibitor in that it harbored an unusual pattern of conserved cysteines. There were two intra-domain disulfide bridges in domain 1 (Cys1-Cys19 and Cys8-Cys35) and three in domain 2 (Cys41-Cys70, Cys45-Cys64 and Cys53-Cys82) (Fig. 1A). Using a BLAST search, the sequence identity between CrSPI-1 and the most homologous member of the Kazal family of inhibitors was analyzed. The observed pattern of S-S bridges in CrSPI-1 was more similar to the non-classical Kazal-type group 1 inhibitors from sea anemone and crayfish than to the mammalian and avian inhibitors (Figs. 2 and S2).

A search for topologically similar proteins within the PDB database was performed with the DALI program [37]. There were no structural homologs of full-length CrSPI-1; however, the DALI search with individual domains showed similarity with several Kazal-type inhibitors. The DALI search with domain 1 (residues Cys1-Glu40) against the pdb database showed that it was structurally similar to leech-derived tryptase inhibitor (pdb code 1ldt) with an rmsd of 2.0 Å for 38 Cα atoms (30% sequence identity). This was followed by dipetalin, a Kazal-type thrombin inhibitor (rmsd = 2.0 Å for 37 Cα atoms; 27.5% identity; pdb code 1bq) and leech-derived tryptase inhibitor (pdb code 1an1), which has an rmsd of 2.4 Å for 29 Cα atoms with 20% sequence identity. Most of the structurally characterized Kazal-type inhibitors contain a single domain, except for rhodniin. Rhodniin is the only structurally characterized Kazal-type inhibitor with two domains in complex with thrombin (1:1 ratio). Both CrSPI-1 domains show significant structural similarity to the respective domains of rhodniin. However, full-length CrSPI-1 could not be superimposed with full-length rhodniin. This indicates that the relative orientations of the two domains are different. Notably, Rhodniin binds with only a single serine protease (thrombin) whereas CrSPI-1 binds with two serine protease molecules. This might be due to the variations of the reactive site loop residues at the prime side of the substrate binding site, which might dictate the specificity of these inhibitors towards a particular serine protease.

The comparison of CrSPI-1 domains 1 and 2 revealed that the core regions superimposed with an rmsd of 3.9 Å for 24 Cα atoms (Fig. 3). The sequence alignment showed 27% identical (42% similar) residues between the two domains, including the highly conserved cysteines (Fig. 2). The highly conserved S-S bridges probably maintain the architecture of these domains. The overall fold similarity between these two domains, together with the structural homology with other Kazal-type inhibitors, suggests that
all Kazal-type inhibitors evolved from a common ancestral gene via duplication while maintaining divergent amino acid sequences. A study by Merckel et al. [38] showed that the similarities between the tertiary structures are indicators of gene duplication.

CrSPI-1: subtilisin complex

The CrSPI-1: subtilisin complex is a heterotrimer in solution, which was confirmed by gel filtration, non-reducing SDS-PAGE [25] and Isothermal Titration Calorimetry (ITC) experiments (Fig. S3). This is consistent with the observation of a heterotrimer of a 1:2 ratio of CrSPI-1 to subtilisin in the asymmetric unit of the crystal (Fig. 1). Both domains of CrSPI-1 act as heads that bind with two individual subtilisin molecules. Our earlier work showed that domain 1 alone did not react with subtilisin, whereas domain 2 showed high affinity for subtilisin with a Ki of 2.6 nM [9]. However, the structural studies revealed that although both domains indeed bind to subtilisin, only domain 2 showed tight interactions with subtilisin. We hypothesize some possible reasons for the tighter interaction of domain 2 in the section “Rigidity of the RSL.”

CrSPI-1 reactive site loop interactions with subtilisin

The reactive site loop (RSL) of domain 1 of CrSPI-1 binds at the active site region of subtilisin from P3 (Cys1) to P3′ (Lys6) in a substrate-like manner (Fig. 4A, panel 1). There were 11 hydrogen bonding contacts between the domain 1 RSL and subtilisin. Of these, five hydrogen bonding contacts were mediated by side chains. The main chain amide group of the P1 residue His3 was involved in a hydrogen bonding contact with O' of the catalytic Ser220 (2.63 Å). The carbonyl O atom of the P1...
His3 interacted with the backbone NH group of the active site Ser220 and N\textsuperscript{\text{\textbeta}} of Asn154 of subtilisin (Table S1). In Kazal-type inhibitor complexes, the P1 residue alone makes approximately 50% of the hydrogen bonding interactions with the active site residues [42,43]. In the CrSPI-1: subtilisin complex, the P1 residue His3 of the CrSPI-domain 1 contributes 5 hydrogen bonding contacts (or 45% of the total hydrogen bonding interactions) with subtilisin. Pro2 occupies the P2 position and makes hydrophobic interactions with subtilisin. In most of the canonical serine protease inhibitors, the P3 residue is engaged in a disulfide bond. In CrSPI-1, Cys1 is in the P3 position and engaged in a disulfide bond with Cys19, which is part of the hydrophobic core that consists of Leu125 (subtilisin) and Pro2 (CrSPI). Thr4 is in the P1’ position, and Tyr5 and Lys6 are in P2’ and P3’ positions, respectively. The P2’ residue Tyr5 maintains stacking interactions with Phe188 of subtilisin (Fig. 4A, panel 1).

In domain 2 of CrSPI-1, Glu49 takes the position of His3 in the P1 pocket of domain 1 and makes 8 (73%) of the 11 total hydrogen bonding contacts with subtilisin. The remaining hydrogen bonding contacts with subtilisin are from P1’ Glu49, P2’ Tyr50 and P4 Val45 of domain 2 (Table S2). Similar to the P2’ residue Tyr5 of domain 1, Tyr50 takes the P2’ position in domain 2 and maintains stacking interactions with Phe188 of subtilisin (Fig. 4A, panel 2). The reactive site residues are well defined in the electron density map (Fig. 4C). In addition, the P3 residue Cys46 mediates hydrophobic interactions with the side chain of Leu125 of subtilisin. Overall, the RSL interactions in domain 2 clearly indicate that the P1 residue Glu48 of domain 2 is the primary mediator of the interaction with subtilisin. These observations show that domain 1 may not be a strong inhibitor of subtilisin compared to domain 2. Notably, domain 2 inhibits subtilisin with a Ki of 2.5 nM, whereas CrSPI-1 domain 1 did not inhibit subtilisin [9]. However, the Ki for full-length CrSPI-1 was 1.43 nM.

Rigidity of the RSL
The conformation adopted by the reactive site loops (RSLs) of both domains of cCrSPI were similar and superimposable with an rmsd of 0.80 Å for the residues from positions P3 to P3’ (Fig. 4B). The RSLs in several families of serine protease inhibitors in many complexes and in different crystal environments adopt similar conformations [Fig. S2]. Similar to other members of the Kazal family of inhibitors, the disulfide bonds formed by cysteine...
residues at the P3 and P5’ positions (Cys1 and Cys8 in domain 1 and Cys46 and Cys54 in domain 2) of CrSPI-1 may hold the RSL in a relatively rigid conformation.

The backbone torsion angles ($\psi/\phi$ angles) of the RSLs of CrSPI-1 were similar to each other and to other protease inhibitors, such as OMTKY3 and Eglin-C complexed with subtilisin (Table S3). In addition, the torsion angles of the RSL backbones of CrSPI-1 were similar to several structurally unrelated serine protease inhibitor complexes and confirm the canonical binding mode for CrSPI-1 (Table S3).

Superimposition of the active site region of the subtilisin complexes of OMTKY3 and Eglin C with the CrSPI-1: subtilisin complex revealed that different inhibitor loop sequences can be accommodated within the substrate binding site of subtilisin with minimal side chain rearrangement (Fig. S2). Furthermore, it indicates that the RSLs of these inhibitors have a rigid conformation, which is supported by the fact that both main chain and side chains of subtilisin remain in the same relative conformation in several complexes and in the uncomplexed form. In domain 2, two out of three S-S bridges (Cys41-Cys70 and Cys45-Cys64) are connected to the central $\alpha$-helix region and maintain a rigid conformation for the RSL of this domain, unlike the RSL in domain 1. Notably, the presence of three disulfide bonds in domain 2 explains the rigidity, specificity and potency of this domain towards its particular cognate protease subtilisin. It is worth mentioning here that in the case of OMSVP3, engineering an additional disulfide bridge increased the specificity for only one protease, $\alpha$-chymotrypsin, and decreased the specificity for other proteases [23].

In addition to the S-S bridges, there are three important internal hydrogen bonding contacts that maintain the rigidity of the RSLs in CrSPI-1. These contacts in domain 1 are (1) the P2-P1’ hydrogen bond between the carbonyl oxygen of Pro2 and amide nitrogen of Thr4 of the reactive site loop, (2) the hydrogen bond between Asn18 and Phe21, and (3) the interaction of N92 of Asn18 with the main chain carbonyl atoms of Pro2 and Thr4 at the P2 and P1’ positions of the RSL. Similarly, in domain 2, the interactions through the side chains of Asn62, Thr47 and Glu49 maintain the rigidity of the RSL. The interactions between Thr47 and Glu49 are similar to the interactions of P2-P1’ of domain 1. In order to reduce the entropic cost of binding, the RSLs of Kazal-type inhibitors and of CrSPI-1 in particular are firmly held in a preferred conformation that is complementary to the substrate-binding site of specific proteases. A rigid conformation of the RSL is thought to prevent proteolytic cleavage of the inhibitor upon interaction with proteases [43].

In order to verify the hypothesis that the rigidity of the RSL protects against the proteolytic cleavage by subtilisin, we investigated the interaction of a peptide (VCTEEY) derived from the RSL region of domain 2 of CrSPI-1 using Isothermal Titration Calorimetric (ITC) experiments (Fig. 5). Titration with the VCTEEY peptide did not resemble a hydrolytic process but a hydrolytic process. Nevertheless, the titration of VCTEEY showed saturation characteristics and was analyzed as a binding process that exhibited an enthalpically driven event with an affinity ($K_a$) of $2 \times 10^5$ M$^{-1}$ ($K_t = 500 \mu$M) (Fig. 5). This indicates that the rigidity of the RSL peptide is not the only factor that affects the proteolytic cleavage of the inhibitor, but the sequence of the interacting peptide might also be important for inhibition. The rigidity and compactness of the domains of CrSPI-1 are due to the presence of disulphide bridges. Hence no structural changes are anticipated between native and recombinant forms. Moreover, recently we have determined the structure of domain 1 (pdb code 3PIS) which superimposed well with the CrSPI-1 domain 1 of the complex and observed no conformational changes [57].

Discussion

Specificity of the CrSPI-1 domains

Kazal-type serine protease inhibitors are single or multi-domain proteins with domains that usually have different specificities towards a particular protease [44]. Although the sequence of the RSL tends to be highly variable, the specificity of the overall molecule for a protease is dictated by the P1 residue. His3 is the P1 residue in the RSL of domain 1, whereas Glu46 is the P1 residue in the RSL of domain 2. Several Kazal-type inhibitors with Glu in the P1 site are known to inhibit subtilisin. For instance, the five-domain
shrimp Kazal inhibitor, SPIPm2, has two subtilisin inhibiting domains with P1 Glu residues [45], and EPI1, a Kazal-like Protease Inhibitor from Phytophthora infestans [46], has Glu as the P1 residue. Thus, the presence of Glu in the P1 position makes the inhibitor more specific for subtilisin [43,47]. Glu49 is buried in the S1 pocket with 8 hydrogen bonding contacts with subtilisin compared to His3, the P1 residue of domain 1, which makes only five hydrogen bonding contacts. This suggests that Glu49 in the P1 pocket of domain 2 is more specific and maintains a tight interaction with subtilisin. The RSL residues of domains 1 and 2 contribute 445 Å² and 530 Å² of buried surface area, respectively, which accounts for approximately 76% and 90% of the binding interfaces with subtilisin, respectively. Thus, the RSL of domain 2 contributes a greater buried area compared to the RSL of domain 1. Notably, domain 2 inhibits subtilisin with a Ki of 2.6 nM [9]. Furthermore, the ITC experiments on full-length CrSPI-1 revealed the presence of two non-identical binding sites; each domain binds one subtilisin molecule. These results suggest

Figure 6. Model of Furin-CrSPI-Subtilisin heterotrimer complex. A schematic illustration on how CrSPI elicits a dual-inhibitory activity to regulate serine protease-driven antimicrobial response during acute phase innate immune response while maintaining homeostasis under naïve condition. Essentially, CrSPI employs its domain 1 (purple) to targets host’s serine proteases (e.g. Furin, in brown) and its domain 2 (yellow) to target the microbial serine proteases (e.g. Subtilisin A in blue) to elicit this dual-inhibitory and regulatory mechanism.

doi:10.1371/journal.pone.0018838.g006
that domain 2 of CrSPI-1 most likely binds to a subtilisin molecule first, and then domain 1 binds a second subtilisin at a slightly lower lower affinity.

**CrFurin and CrSPI-1**

Previously we have shown the inhibition of CrFurin by CrSPI-1 in an ex-vivo inhibition assay [9]. However the precise mechanism of this inhibition is not yet established. CrFurin, a homolog of subtilisin, shows specificity towards paired basic residues for cleavage [40]. Furin cleaves a wide range of proproteins at the consensus sequence Arg-Xaa-Lys/Arg-Arg-Arg [41], and the minimal consensus sequence for Furin is Arg-Xaa-Xaa-Arg-Arg [42]. Although neither of the two domains of CrSPI-1 contains this consensus sequence, the domain 1 RSL consists of two basic residues (His3 and Lys6). Notably, Kazal-type inhibitors recognize proteases in a substrate-like manner. Furin might not cleave this RSL due to its rigid conformation and the presence of basic residues probably renders it as a substrate-like inhibitor. Several proteins have been engineered to contain a Furin consensus motif in their reactive site loops to thereby inhibit Furin. For instance, mutation of the turkey ovomucoid third domain RSL sequence from Ala-Cys-Thr-Leu18 to Arg-Cys-Lys-Arg18 or mutation of the α1-antitrypsin Portland RSL from Ala353-Ile-Pro-Met358 to Arg353-Ile-Pro-Arg358 both produced effective inhibitors for Furin [49]. These inhibitors mimic highly specific substrates. Due to the tight binding and specific rigid conformation of the RSL residues, these inhibitors are able to arrest the enzymatic reaction at the intermediate stage of hydrolysis of the peptide bond [50].

**Model of the domain 1 CrSPI-1: Furin complex**

Furin is a subtilisin-related serine protease, and a member of the proprotein convertase family. Although, subtilisin is obtained from bacteria belonging to the *Bacillus* species, Furin is considered as a eukaryotic version of subtilisin, which is ubiquitously expressed [48,51–53]. The Furin binding mode of CrSPI-1 was predicted based on the crystal structures of CrSPI-1: subtilisin and human Furin (Pdb code 1P8J). The host Furin is a homolog of the bacterial subtilisin, and these serine proteases share a sequence identity of 23% (~40% similarity). Moreover, the structure of subtilisin superimposes on the subtilisin-like domain of Furin with an rmsd of 2 Å for 268 Cz atoms of 274 Cz atoms of subtilisin. The domain 1 CrSPI-1: subtilisin complex was superimposed onto the structure of Furin (pdb code 1P8J), and the coordinates of CrSPI-1 were copied to Furin to generate the CrSPI-1:Furin complex model (Figs. 6, S4 and S5). The RSL region from P3 (Cys1) to P3’ (Lys6) occupies the substrate binding site of human Furin. This is expected as Furins and subtilisins have highly homologous structures. However, this needs to be experimentally verified. Based on our analysis, we speculate that domain 1 of CrSPI-1 is responsible for the inhibition of CrFurin. Moreover domain 2 is shown to be an inhibitor of Subtilisin [9]. It is possible that the two domains of CrSPI-1 are products of a gene duplication event that generated a dual specificity inhibitor with one of the domains functioning as an inhibitor of host proprotein-converting subtilisin-like enzyme, CrFurin, and another domain functioning as a pathogen-specific protease (subtilisin) inhibitor.

**Possible immunomodulatory activities of CrSPI-1**

The preference of CrSPI-1 domain 2 for the microbial protease subtilisin suggests that this isoform targets serine proteases of invading microbes. In fact, subtilisin is the virulence factor used by the pathogen to gain entry into host cells during an infection. Besides direct suppression of the microbial proteases, CrSPI-1 has been shown to target the host’s endogenous proteases like the CrFurin. This is supported by the observation that in response to an infection, a dynamic reciprocal expression profile of the CrSPI-1 and CrFurin transcripts was observed [9]. CrSPI-1 has been proposed to regulate the serine protease-driven antimicrobial defense during the acute phase of infection while maintaining homeostasis under normal physiological conditions [9]. It has also been found to interact with microbial subtilisins as well as an endogenous C3 complement from the host, which is required for the clearance of pathogens. It is possible that CrSPI-1 serves an immunomodulatory function by interacting with host and microbial proteases through its independent domains (Fig. 6). We hypothesize that the two domains of CrSPI-1 have different substrate preferences. Domain 2 might be more potent and specific towards the bacterial protease, subtilisin, whereas domain 1 likely interacts with the host protease Furin. The structure of the CrSPI-1: subtilisin ternary complex is the first step towards an understanding of the molecular perspective of the antibacterial response while maintaining homeostasis for the host through such a dual-inhibitory mechanism. The involvement of serine proteases as the virulent factor in the entry of several microbial pathogens makes it an important target for the treatment of several infectious diseases.

**Supporting Information**

**Figure S1** Alignment of amino acid sequences of non-classical group I Kazal-type inhibitors AEI, CrSPI-1 domain I, domain II and a selected classical Kazal-type inhibitor OMSVP3. The sequences were aligned using CLUSTAL-W. The reactive site is denoted with an arrow. Disulfide bonds are linked as follows: α-β, II-IV, and III-VI for the non classical group I inhibitors and I-V, II-IV, and III-VI for the classical inhibitors. In nonclassical group II inhibitor family, there is an additional disulfide bridge between α and β half cystines. (TIF)

**Figure S2** Conformations of the reactive site loop (RSL). Superimposition of the reactive site loops of domain-1 (magenta), domain-2 (yellow), Eglin C (gray) and OMTKY3 (cyan). The RSLs are shown in stick representation whereas the substrate binding site of subtilisin is shown in surface representation. These figures were generated by using PyMol. (TIF)

**Figure S3** Isothermal Titration Calorimetric (ITC) curve for rCrSPI-1 titrated against subtilisin at 37°C. Each peak represents the injection of rCrSPI-1 0.2 mM into the ITC cell containing subtilisin 0.012 mM, in buffer PBS pH 7.4, 10 mM BME. A sequence of 18 injections, each injection consisting of 2 μL of ligand solution, was performed. The experimental data were fitted considering a model in which CrSPI-1 binds two Subtilisin molecules, either employing a general model based on the overall association parameters or considering two non-identical and independent binding sites in CrSPI-1 (32). Binding association constants of 2.4x10^6 M^-1 and 1.7x10^4 M^-1 were obtained from non-linear regression analysis, corresponding to dissociation constants of 0.42 and 59 μM, respectively. (TIF)

**Figure S4** Cz trace for the heterotrimer Furin-CrSPI-Subtilisin complex model. Furin and subtilisin share a sequence identity of 23%. The Furin:CrSPI-1 complex model was generated by superimposing domain-1 CrSPI-1: subtilisin complex onto the structure of Furin (pdb code: 1p8j), which yielded an rmsd of 2Å for
268 Cα out of 274 Cα atoms of subtilisin. The Furin-CrSPI-1-Subtilisin heterotrimer complex was generated using the modeled Furin-CrSPI-1-domain-1 and subtilisin-CrSPI-1-domain-2 complex crystal structure.

**Figure S5** Surface representation for Furin and Subtilisin, and backbone trace representation for CrSPI-1 of the heterotrimer model.

**Table S1**

| References |
|------------|
| 1. Jiravanichpaisal P, Lee BL, Soderhall K (2006) Cell-mediated immunity in arthropods: hematopoiesis, coagulation, melanization and opsonization. Immunobiology 211: 213–36. |
| 2. Theopold U, Schmidt O, Soderhall K, Dushay MS (2004) Coagulation in arthropods: defence, wound closure and healing. Trends Immunol 25: 289–94. |
| 3. Armstrong PB (2001) The contribution of protease inhibitors to immune defense. Trends Immunol 22: 47–52. |
| 4. Ding JL, Navas MA, Srd, Ho B (1993) Two forms of factor C from the anemocytes of Carcinopus rotundicauda. Purification and characterization. Biochem Biophys Acta 1202: 149–156. |
| 5. Cerenius L, Soderhall K (2004) The prophenoloxidase-activating system in invertebrates. Immuno Rev 198: 116–26. |
| 6. Ding JL, Ho B (2001) New era in pyrogen testing. Trends Biotechnol 19: 277–81. |
| 7. Iwanaga S, Kawabata S (1998) Evolution and phylogeny of defense molecules associated with innate immunity in horseshoe crab. Front Biosci 3: D973-84. |
| 8. Kanout MR (1999) Serine protease inhibitors in arthropod immune system. Dev Comp Immunol 23: 291–301. |
| 9. Jiang N, Thangamani S, Chor CF, Wang SY, Winarsih I, et al. (2009) A Novel Serine Protease Inhibitor Acts as an Immunomodulatory Switch while Maintaining Homeostasis. J Innate Immun 1: 465–479. |
| 10. Ding JL, Wang LJ, Ho B (2004) Current genome-wide analysis on serine proteases in innate immunity. Curr Genomic 5: 147–155. |
| 11. Hiemstra PS (2002) Novel roles of protease inhibitors in infection and inflammation. Biochem Soc Trans 30: 116–120. |
| 12. Chavanas S, Bodemer C, Rochat A, Hamel-Teillac D, Ali M, et al. (2000) A Four-Domain Kazal Proteinase Inhibitor from Crayfish Blood Cells. J Biol Chem 275: 20227–20234. |
| 13. Muta T, Iwanaga S (1996) The role of hemolymph coagulation in innate defense. Trends Immunol 22: 47–52. |
| 14. Bode W, Huber R (1992) Natural protein proteinase inhibitors and their crystal structures. Biochim Biophys Acta 117: 159–171. |
| 15. Holm L, Sander C (1991) Database algorithm for generating protein backbone structures by the maximum-likelihood method. Acta Crystallogr A 47: 110–9. |
| 16. Stoll VS, Eger BT, Hynes RC, Martichonok V, Jones Jr, et al. (1998) Differences in binding modes of enantioomers of 1-acetamido boronic acid based protease inhibitors: crystal structures of gamma-chymotrypsin and subtilisin Carlsberg complexes. Biochemistry 37: 451–462. |
| 17. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66: 12–21. |
| 18. Schmitke JL, Strem LJ, Kishanov AV (1998) Comparison of x-ray crystal side-chain co-ordinates from a C α trace application to model building and side-chain co-ordinates from a C α trace application to model building and detection of co-ordinate errors. J Mol Biol 281: 833–843. |
| 19. PLoS ONE | www.plosone.org 10 April 2011 | Volume 6 | Issue 4 | e18838 |

**Acknowledgments**

The authors acknowledge the X299-A beamline, National Synchrotron Light Source, and Brookhaven National Laboratory for the data collection, and thank Dr Anand Saxena for help during the data collection. We thank A/P K.Swaminathan for useful suggestions.

**Protein Data Bank code**

Coordinates of CrSPI-1: subtilisin complex have been deposited in the Protein Data Bank [http://www.pdb.org] with accession code 3QTL.

**Author Contributions**

Conceived and designed the experiments: JS, JLD ST. Performed the experiments: RTS ST. Analyzed the data: JS RTS AV C JS. Contributed reagents/materials/analysis tools: JLD BH AVC JS. Wrote the paper: RTS ST AV C JLD JS.

**References**

1. Jiravanichpaisal P, Lee BL, Soderhall K (2006) Cell-mediated immunity in arthropods: hematopoiesis, coagulation, melanization and opsonization. Immunobiology 211: 213–36.
2. Theopold U, Schmidt O, Soderhall K, Dushay MS (2004) Coagulation in arthropods: defence, wound closure and healing. Trends Immunol 25: 289–94.
3. Armstrong PB (2001) The contribution of protease inhibitors to immune defense. Trends Immunol 22: 47–52.
4. Ding JL, Navas MA, Srd, Ho B (1993) Two forms of factor C from the anemocytes of Carcinopus rotundicauda. Purification and characterization. Biochem Biophys Acta 1202: 149–156.
5. Cerenius L, Soderhall K (2004) The prophenoloxidase-activating system in invertebrates. Immuno Rev 198: 116–26.
6. Ding JL, Ho B (2001) New era in pyrogen testing. Trends Biotechnol 19: 277–81.
7. Iwanaga S, Kawabata S (1998) Evolution and phylogeny of defense molecules associated with innate immunity in horseshoe crab. Front Biosci 3: D973-84.
8. Kanout MR (1999) Serine protease inhibitors in arthropod immune system. Dev Comp Immunol 23: 291–301.
9. Jiang N, Thangamani S, Chor CF, Wang SY, Winarsih I, et al. (2009) A Novel Serine Protease Inhibitor Acts as an Immunomodulatory Switch while Maintaining Homeostasis. J Innate Immun 1: 465–479.
10. Ding JL, Wang LJ, Ho B (2004) Current genome-wide analysis on serine proteases in innate immunity. Curr Genomic 5: 147–155.
11. Hiemstra PS (2002) Novel roles of protease inhibitors in infection and inflammation. Biochem Soc Trans 30: 116–120.
12. Chavanas S, Bodemer C, Rochat A, Hamel-Teillac D, Ali M, et al. (2000) Mutations in SPINK5, encoding a serine protease inhibitor, cause Netherton syndrome. Nat Genet 25: 141–2.
13. Muta T, Iwanaga S (1996) The role of hemolymph coagulation in innate immunity. Curr Opin Immunol 8: 41–7.
14. Bode W, Huber R (1992) Natural protein protease inhibitors and their interaction with proteases. Eur J Biochem 204: 433–451.
15. Hemmi H, Kumaizaki T, Yoshizawa-Kumagay K, Nishitachi Y, Yoshida T, et al. (2005) Structural and functional study of an Anemone elastase inhibitor, a “nonclassical” Kazal-type inhibitor from Anemonia sulcata. Biochemistry 44: 9626–36.
16. Tschersche H, Kolkenbrock H, Bode W (1987) The covalent structure of the elastase inhibitor from Anemona sulcata - a “non-classical” Kazal-type protein. Biol Chem Hoppe Seyler 368: 1297–304.
17. Johansson MW, Keyser P, Soderhall K (1994) Purification and cDNA cloning of a four-domain Kazal protease inhibitor from crayfish blood cells. Eur J Biochem 223: 389–94.
18. Odum L, Bundgaard JR, Johnsen AH (1999) A Kazal-type trypsin inhibitor from the protochordate Ciona intestinalis. Eur J Biochem 259: 207–15.
19. Nirmala X, Kodrik D, Zureve M, Sefula F (2001) Insect silk contains both a Kunzite type and a unique Kazal-type protease inhibitor. Eur J Biochem 269: 2604–73.
20. Sommerhoff CP, Sollner C, Mensle R, Pechotka GP, Auerwald EA, et al. (1994) A Kazal-type inhibitor of human mast cell tryptase: Isolation from the medical leech Hirudo medicinalis, characterization, and sequence analysis. Biochim Biophys Acta 1202: 375–835.
21. Fink E, Rehm H, Gippenner C, Bode W, Edlitz M, et al. (1986) The primary structure of bleduin B from the leech Hirudo medicinalis. Biol Chem Hoppe-Seyler 357: 653–654.
22. Friedrich T, Kroger B, Bialojan S, Lemaire HG, Hofflun HW, et al. (1993) A Kazaltype inhibitor with thrombin specificity from Rhabdias prolixus. J Biol Chem 268: 16216–16222.
42. Kleantous C (2000) Protein-protein recognition. Oxford University Press. 315 p.
43. Maynes, JT, Cherney MM, Qasim MA, Laskowski M, Jr., James MN (2005) Structure of the subtilisin Carlsberg-OMTKY3 complex reveals two different ovomucoid conformations. Acta Crystallogr D Biol Crystallogr 61: 580–8.
44. Dompuls S, Tassanakajon A, Rimpahanitchayakit V (2009) Domain inhibitory and bacteriostatic activities of the five-domain Kazal-type serine proteinase inhibitor from black tiger shrimp Penaeus monodon. Dev Comp Immunol 33: 481–8.
45. Ohum L, Bundgaard JR, Johnsen AH (1999) A Kazal type trypsin inhibitor from the protochordate Ciona intestinalis. Eur J Biochem 259: 872–876.
46. Tian M, Kamoun S (2005) A two disulfide bridge Kazal domain from Phytophthora exhibits stable inhibitory activity against serine proteinases of the subtilisin family. BMC Biochem 6: 15.
47. Somprasong N, Rimpahanitchayakit V, Tassanakajon A (2006) A five-domain Kazal-type serine proteinase inhibitor from black tiger shrimp Penaeus monodon and its inhibitory activities. Dev Comp Immunol 30: 998–1008.
48. Thomas G (2002) Furin at the cutting edge: from protein traffic to embryogenesis and disease. Nat Rev Mol Cell Biol 3: 753–60.
49. Lu W, Zhang W, Molloy SS, Thomas G, Ryan K, Chiang Y, et al. (1993) Arg15-Lys17-Arg18 turkey ovomucoid third domain inhibits human furin. J Biol Chem 268: 14307–14305.
50. Radisky ES, Koshland DE, Jr. (2002) A clogged gutter mechanism for protease inhibitors. Proc Natl Acad Sci U S A 99: 10316–21.
51. Nakayama K (1997) Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. Biochem J 327: 625–635.
52. Rockwell NC, Krysan DJ, Komiyama T, Fuller RS (2002) Precursor processing by kex2/furin proteases. Chem Rev 102: 4525–48.
53. Scamuffa N, Calvo F, Chretien M, Seidah NG, Kharib AM (2006) Proprotein convertases: lessons from knockouts. FASEB J 20: 1954–63.
54. DeLano WL (2002) The PyMOL Molecular Graphics System DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.
55. Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Cryst 24: 946–950.
56. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: multiple sequence alignments in PostScript. Bioinformatics 15: 305–8.
57. Giri PK, Tang X, Thangamani S, Shenoy RT, Ding JL, et al. (2010) Modifying the substrate specificity of Carcinoscorpius rotundicauda serine protease inhibitor domain 1 to target thrombin. PLoS ONE 5(12): e15258. Doi:10.1371/journal.pone.0015258.