Tri-domain Bifunctional Inhibitor of Metallocarboxypeptidases A and Serine Proteases Isolated from Marine Annellid 

This study describes a novel bifunctional metallocarboxypeptidase and serine protease inhibitor (SmCI) isolated from the tentacle crown of the annelid Sabellastarte magnifica. SmCI is a 165-residue glycoprotein with a molecular mass of 19.69 kDa (mass spectrometry) and 18 cysteine residues forming nine disulfide bonds. Its cDNA was cloned and sequenced by RT-PCR and nested PCR using degenerated oligonucleotides. Employing this information along with data derived from automatic Edman degradation of peptide fragments, the SmCI sequence was fully characterized, indicating the presence of three bovine pancreatic trypsin inhibitor/Kunitz domains and its high homology with other Kunitz serine protease inhibitors. Enzyme kinetics and structural analyses revealed SmCI to be an inhibitor of human and bovine pancreatic metallo- and serine-type proteases. SmCI is also capable of inhibiting bovine pancreatic trypsin, chymotrypsin, and porcine pancreatic elastase in varying measures. When the inhibitor and its nonglycosylated form (SmCI N23A mutant) were overproduced recombinantly in a Pichia pastoris system, they displayed the dual inhibitory properties of the natural form. Similarly, two bi-domain forms of the inhibitor (recombinant rSmCI D1-D2 and rSmCI D2-D3) as well as its C-terminal domain (rSmCI-D3) were also overproduced. Of these fragments, only the rSmCI D1-D2 bi-domain retained inhibition of metallocarboxypeptidase A but only partially, indicating that the whole tri-domain structure is required for such capability in full. SmCI is the first proteinaceous inhibitor of metallocarboxypeptidases able to act as well on another mechanistic class of proteases (serine-type) and is the first of this kind identified in nature.

Metallocarboxypeptidases (CPs) are an important class of enzymes that catalyze the hydrolysis of peptide bonds at the C terminus of peptides and proteins. Besides a role in digestive protein degradation, these enzymes are also key elements of selective proteolysis-regulated physiological processes such as blood coagulation/fibrinolysis, inflammation, prohormone and neuropeptide processing, local anaphylaxis, and insect/plant-attack/defense strategies, among others (1, 2). The biological actions of many proteases are controlled by their interaction with specific proteinaceous inhibitors. Unlike endoproteases, for which numerous examples of protein inhibitors have been reported, naturally occurring metallocarboxypeptidase inhibitors are somewhat limited, and so far, they have only been identified in Solanacea, tomato, and potato (PCI) (3–5), the intestinal parasite Ascaris suum (ACI, Ascaris carboxypeptidase inhibitor) (6, 7), the medicinal leech Hirudo medicinalis (LCI) (8), the ticks Rhipicephalus bursa (TCI) (9) and Hemaphysalis longicornis (H1TCI) (10), and in rat and human tissues (latexin or endogenous carboxypeptidase inhibitor (ECI)) (11, 12).

The mechanisms of the inhibitory actions of PCI, LCI, and TCI on CPs rely upon interaction of their C-terminal tail with
the active site cleft of the enzyme in a manner that mimics substrate binding (2, 5, 8, 13). Additionally, TCI anchors to the surface of CPs of the A/B-type in a double-headed manner not observed for the other protein inhibitors (13). However, the C-tail of mammalian tissue protein inhibitors does not seem to be a suitable substrate for CPs; such proteins interact with the enzymes through one loop located at the interface of their two subdomains (14). In addition, the pro-regions of procarboxypeptidases, which fold as independent globular domains, position their internal inhibitory loop on the active site cleft of the enzyme rendering the enzyme inhibited (15, 16). All these inhibitors are specific for the A/B metallo-CP subfamily, regardless of their substrate preferences (2).

In general, protein inhibitors of proteases belonging to different mechanistic classes are uncommon. Such inhibitors may feature one or more inhibitory domains, such as SHPI-1 (with one BPTI/Kunitz-type domain), which is able to inhibit serine, cysteine, and aspartic proteases (17), or equistatin (with three thyroglobulin-1 domains), which inhibits cysteine and aspartic proteases (18, 19), among others. However, there have been no descriptions to date either of a multifunctional inhibitor able to inhibit CPs and proteases belonging to distinct mechanistic classes nor of inhibitors of CPs with a typical Kunitz structure with the capacity to inhibit several serine proteases.

Among the available natural sources of protease inhibitors, one of the most attractive and rather unexplored is the marine fauna, especially invertebrates (including numerous phyla, genera, and species). Several such inhibitors capable of independently inhibiting proteases of different mechanistic classes have been described, particularly in the phyla Cnidaria (17–27), Mollusca (28–31), and Annelida (32, 33). However, no inhibitors of CPs of this type had been described in such sources until now. No inhibitory activity against CPs nor of inhibitors of CPs with a typical Kunitz structure were synthesized by TIB Biomol (Germany); the Nucleospin and NucleoTrap kits were both from Clontech. TOP10F’ cells, pPICZaA vector, Pichia EasyComp transformation kit, and Zeocin™ were supplied by Invitrogen; and pGEM-T-easy vector was supplied by Promega.

**Purification of SmCI**—The marine invertebrate was collected at north of Havana, Cuba, and was taxonomically identified by specialists of the Cuban National Institute of Oceanology. The tentacle, or feathered, crowns of the animals were separated from the body, homogenized (2:1 v/w), and centrifuged. The supernatant was clarified by heating at 60 °C for 20 min and centrifuged. The heated extract was loaded in three steps onto a CPA-glyoxy-agarose column (0.9 × 5.5 cm) prepared as 1.3 mg of immobilized CPA per ml of gel according to the general procedure described for other enzymes (35), with some modifications. Unbound proteins were eluted by washing the column with a sufficient quantity of equilibration buffer (0.05 M Tris-HCl, 0.5 M NaCl, 10⁻⁵ M ZnCl₂, pH 7.0). Proteins with CPA inhibitory activity were eluted by increasing the pH to 10.4 through the addition of 0.05 M glycine, 0.04 M NaOH buffer at a linear flow rate of 24 cm/h. Inhibitor-containing fractions were lyophilized and subjected to reverse phase HPLC (Vydac C8 column). The homogeneity of the purified inhibitor was verified by SDS-PAGE (36) and MALDI-TOF mass spectrometry. The SDS-polyacrylamide gel was stained with Coomassie Blue R-250. Prestained molecular weight standards were used.

**Inhibitory Activity**—Inhibition assays were performed by preincubating the inhibitor with the enzymes for 10 min at room temperature before adding the substrate. The inhibitory activities of CPA (EC 3.4.17.1) and carboxypeptidase B (CPB; EC 3.4.17.2) were determined using bovine pancreatic CPA and porcine pancreatic CPB, respectively (both at 1.5 × 10⁻⁸ M) as models and the substrates N-(4-methoxyphenylazoformyl)-L-phenylalanine and N-(4-methoxyphenylazofomyl)-L-lysine (at 0.1 mM in the assays) (37, 38). Hydrolysis of both substrates was followed at 350 nm (37, 38) at 15-s intervals for 10 min at 25 °C (19 × 10⁻⁶ μM⁻¹ cm⁻¹) in a kinetic spectrophotometer (GE Healthcare).

The activities of the following enzymes were assayed on the following substrates: bovine pancreatic trypsin (EC 3.4.21.4; 1.18 × 10⁻⁸ M in the assay) by monitoring esterase activity against BAPA (1.0 mM in the assay) (39); bovine pancreatic chymotrypsin, porcine pancreatic elastase, bovine pancreatic chymotrypsin, porcine pepsin, and p-nitrophenyl-p-guanidinobenzoate were purchased from Sigma. Papain from *Carica papaya* was supplied by Calbiochem; benzoylarginyl-p-nNA (BAPA), N-(4-methoxyphenylazoformyl)-L-phenylalanine, MeOSuc-(L-Ala)₂-L-Pro-(L-Val-pNA, N-(4-methoxyphenylazofomyl)-L-lysin, Leu-Ser-Phe(NO₂)-(Nle-Ala-Leu-OMet, and Suc-Ala-Ala-Pro-Phe-pNA were supplied by BACHEM (Germany). N-Glycosidase F was from Roche Applied Science. The restriction enzymes SacI, Xhol, NotI, and EcoRI, and avian myeloblastosis virus reverse transcriptase were supplied by Roche Applied Science. Oligonucleotides were synthesized by TIB Biomol (Germany); the Nucleospin and NucleoTrap kits were both from Clontech. TOP10F’ cells, pPICZaA vector, Pichia EasyComp transformation kit, and Zeocin™ were supplied by Invitrogen; and pGEM-T-easy vector was supplied by Promega.

**EXPERIMENTAL PROCEDURES**

**Chemicals and Reagents**—All chemicals were reagent grade. HiTrap™ Q-Sepharose FF (1 ml) and STREAMLINE Direct HST matrix were supplied by GE Healthcare, and C18 Sep-Pack cartridge was supplied by Waters. The precast molecular weight standards were from Bio-Rad and Invitrogen; and bovine pancreatic carboxypeptidase A1 and carboxypeptidase B1, bovine pancreatic trypsin, porcine pancreatic elastase, bovine pancreatic chymotrypsin, porcine pepsin, and p-nitrophenyl-p-guanidinobenzoate were purchased from Sigma. Papain from *Carica papaya* was supplied by Calbiochem; benzoylarginyl-p-nNA (BAPA), N-(4-methoxyphenylazoformyl)-L-phenylalanine, MeOSuc-(L-Ala)₂-L-Pro-(L-Val-pNA, N-(4-methoxyphenylazoformyl)-L-lysin, Leu-Ser-Phe(NO₂)-(Nle-Ala-Leu-OMet, and Suc-Ala-Ala-Pro-Phe-pNA were supplied by BACHEM (Germany). N-Glycosidase F was from Roche Applied Science. The restriction enzymes SacI, Xhol, NotI, and EcoRI, and avian myeloblastosis virus reverse transcriptase were supplied by Roche Applied Science. Oligonucleotides were synthesized by TIB Biomol (Germany); the Nucleospin and NucleoTrap kits were both from Clontech. TOP10F’ cells, pPICZaA vector, Pichia EasyComp transformation kit, and Zeocin™ were supplied by Invitrogen; and pGEM-T-easy vector was supplied by Promega.
motrypsin (EC 3.4.21.1; 6.1 × 10⁻⁹ M in the assay) on Suc-Ala-Ala-Pro-Phe-p-nitroanilide as substrate (1 mM in the assay) (40); porcine pancreatic elastase (EC 3.4.21.36; 5.26 × 10⁻⁸ M in the assay) on MeO-Suc-(t-Ala)₂-t-Pro-l-Val-p-nitroanilide as substrate (1 mM in the assay) (41); papain from papaya (EC 3.4.22.2; 3.0 × 10⁻⁷ M in the assay) on BAPA (1 mM in the assay) as substrate (42); and porcine pepsin (EC 3.4.23.1; 3.1 × 10⁻⁷ M in the assay) on Leu-Ser-Phe(NO₂)₂-Nle-Ala-Leu-OMe as substrate (43). All assays were performed at 25 °C for 3 min. The time needed to reach inhibition equilibrium was previously determined for each assay by preincubating SmCI with the enzymes for 1, 5, 10 or 20 min before substrate addition. For each enzyme, the effect of the substrate concentration was also determined using concentrations equivalent to 0.5 K_M, 1 K_M, and 2 K_M.

Active Concentrations of Trypsin and Inhibitor—The active pancreatic trypsin concentration was determined by titration with the standard solution of p-nitrophenyl-p-guanidino benzotate (4.9 × 10⁻⁴ M in the assay) (44). The active inhibitor concentrations were determined by titrating with increasing concentrations of trypsin (from 0.19 × 10⁻² to 1.7 × 10⁻² M in the assays) using a constant concentration of the inhibitor (8.0 × 10⁻⁸ M) and a preincubation time of 10 min, under conditions of E_0/K_I = 100. Residual activity was assayed using 1 mM BAPA, and the active inhibitor concentration was determined at the equivalence point ([E_o] = [I_o]) indicated in a plot of residual activity against enzyme concentration.

Dissociation Constants (K_I)—Inhibition constants (K_I) for the complexes formed by the inhibitors with the different enzymes were determined using a described method for tight binding inhibition (45). Different amounts of inhibitor were preincubated for 10 min with the enzymes under conditions of E_0/K_I = 10. At each concentration, residual activity (v_r) was measured against the specific substrate (using a substrate concentration equivalent to 1 K_M for each enzyme). The experimental points were adjusted to the equation described for tight binding mechanisms (46) by nonlinear fitting using the GraphPad Prism 5.0 package. True K_I values were calculated using the equation: K_I = K_p(app)/[(S_o/K_M) + 1], incorporating the [S_o] and the K_M value for each enzyme.

Protein Concentration—Protein concentrations were determined by the bicinchoninic acid method (47) using the BCA kit and poly(A)^+ RNA was purified using the Nucleospin kit both according to the manufacturer’s instructions. The first strand of cDNA from each enzyme was synthesized using the oligo(dT) primer (5’-CCGGAATTCCTGACAGCAG-3’; R, 5’-GTACCCAAATGACTGACTTACGGGC-3’) and avian myeloblastosis virus reverse transcriptase according to the supplier’s protocols. For cloning the SmCI cDNA, two degenerated oligonucleotides were designed based on its N-terminal sequence, S1(8–17), 5’-GCNGAYTGYYGNCARTGYCANGCATAYAT-3’ (residues 8–17 from the N-terminal sequence), and S2(13–22) (amino acid residues 13–22) 5’-TGYANGCN- TAYATHTCCNARTGTTYTT-3’ (where Y = (C/T), N = (A/C/G/T), R = (A/G), and H = (A/C/T).

For the first PCR round, SmCI cDNA was amplified using oligonucleotide S1(8–17) and the adaptor oligonucleotide R_p. PCR was conducted as 40 cycles each at 94 °C for 1 min, annealing at 48 °C for 1 min, and extension at 72 °C for 2 min. PCR mixtures were 20-fold diluted and re-amplified using the S2(13–22) and R_n nested oligonucleotides. PCR products were separated by electrophoresis on 2% agarose gels. The main nested PCR product was recovered from the gel and cloned into the pGEM-T-easy vector to generate the pGEM-SmCI S1/S2 construct (corresponding to SmCI amino acid residues from 13 to 165).

Nucleotide Sequencing and Computational Analysis of Sequence Data—The pGEM-SmCI S1/S2 clones generated were sequenced using the SP6 and T7 promoter forward and reverse primers in a DNA sequencer (PerkinElmer Life Sciences). Alignments of nucleotides and amino acid residues were carried out using VECTOR NTI as implemented by the VECTOR NTI Suite 9 program of the InforMax 2003 package. PSI-BLAST and PFAM (48, 49) were used to search nonredundant databases. Multiple sequence alignment was performed by combining Clustal X (50), the protein sequences identified by PSI-BLAST, and the disulfide bridge information derived from the resolved crystal structures using DSSP (51).

Deglycosylation Assay—Samples were deglycosylated with N-glycosidase F. SmCI prepared at 1 mg/ml in 5 mM Tris-HCl buffer, pH 8.0, was incubated with an appropriate volume of N-glycosidase F (1 unit/μl) to achieve a final ratio of 100:1 v/v.
The reaction was left to run for 24 h at 37 °C. The deglycosylation process was monitored by MALDI-TOF MS.

**Heterologous Expression of the Recombinant Inhibitors—**

rSmCI, rSmCI N23A (a nonglycosylated form), the two bi-domains rSmCI D1-D2 and rSmCI D2-D3, and its third domain (rSmCI-D3) were overexpressed in the P. pastoris system. First, the SmCI encoding sequence was completed by PCR using two specific sense gene primers (S3(6–18) and S4(1–9)). These primers were designed from the N-terminal sequence of the protein, S3(6–18) (residues 6–18) 5'–TTGCCAGCTGATAGGAGTGTAATGTACGGCCACTACCTC-3' and S4(1–9) (residues 1–9) 5'TCTCTCGAGAAAACTTTGTTGTGTATGTTCGCCAGCTGAT-3'. As the antisense primer we used S5(163–165), which corresponds to the SmCI C-terminal sequence (residues 163–165) 5'–CCTTCCGCGCGGCTAGCAAGCATT-3'. These primers included the restriction sites for XhoI and NotI (underlined) to clone the gene into the pPICZαA vector, and the Kex 2 cleavage site (sense primer) and stop codon (antisense primer) are shown in boldface type. PCR was conducted under the same conditions as described above using pGEM-SmCI S1/S2 vector as template. SmCI gene was removed from pGEM-T easy vector using XhoI and NotI restriction sites and cloned into pPICZαA vector. The new generated vector, designated pPICZαA-SmCI, was transformed in the KM71H strain (Mut phenotype) of P. pastoris following the manufacturer's protocol (Invitrogen). Several colonies on the YPDS agar/Zeo-supernatants obtained after fermentation were loaded onto an Streamline Direct HST column (2.0 × 25.0 cm) equilibrated with 100 mM sodium citrate, pH 4.0 (buffer A1), using an AKTAprime system (GE Healthcare). Unbound proteins were eliminated by washing the column with 20 mM sodium citrate, pH 4.0 (buffer A2). The recombinant protein was eluted using a gradient of 25–100% buffer B (150 mM Tris-HCl, pH 8.0) in 20 column volumes at 10 ml/min. Fractions showing CPA and/or trypsin inhibitory activity were applied to an anion exchange column of HiTrapTM Q-Sepharose FF (1 ml) (General Electric), previously equilibrated with 20 mM Tris-HCl, pH 8.0, buffer at a flow rate of 1 ml/min. After washing with the equilibrium buffer (6 column volumes), bound inhibitor was eluted with a linear gradient from 0 to 45% of 0.5 M sodium chloride in 20 mM Tris-HCl, pH 8.5, in 20 column volumes at the same flow rate. rSmCI-D3 was purified by combining the reverse phase in the Sep-Pak C18 cartridge with the anion exchange step in Q-Sepharose.

The molecular weights, N-terminal sequences, and inhibitory activity against CPA, CPB, trypsin, and pancreatic elastase of the recombinant forms were assessed as described above for the natural protein. $K_i$ values of the inhibitors against different enzymes were determined using the same kinetic strategy and substrates employed for the natural inhibitor. For the pancreatic elastase inhibitory assay, N-Suc-(Ala)$_3$-pNA was used as substrate in case of rSmCI.

**RESULTS**

**Isolation and Purification of SmCI—**

The proteinaceous inhibitor was isolated from the tentacle crown of S. magnifica by both affinity and reverse phase chromatography. The inhibitory activity recovered after heat treatment at 60 °C (20 min) was 92% indicating the stability of SmCI at this temperature. The heated extract was loaded in three steps onto an affinity column (CPA-glyoxyl-Sepharose). This procedure allows the application of a larger number of inhibitor than standard methods (one application) (Fig. 1A). The inhibitor was strongly bound to the immobilized enzyme and eluted by increasing the pH of the buffer to 10.4. The yield, based on the inhibitory capability on CPA, was 113%, and purification was 102-fold with respect to the crude extract. The main component of the affinity fraction applied to the reversed phase HPLC was purified to homogeneity in appropriate conditions for its molecular and kinetic characterization. SDS-PAGE of the purified material showed a single protein band around 20 kDa (Fig. 1B). This result is consistent with a molecular mass of 19.69 kDa determined by MALDI-TOF MS (Fig. 1C).

**Inhibitory Activity and Selectivity of SmCI—**

Purified SmCI was able to inhibit proteases pertaining to two different mechanistic classes, metallocarboxypeptidases (A subtype, but not B) and serine proteases (trypsin, pancreatic elastase and chymotrypsin). In contrast, using up to a 10-fold molar excess of inhibitor, no blocking action was detected on the activities of papain or pepsin, two standard models for cysteine and aspartic protease.

Equilibrium dissociation constants ($K_i$) for the interaction of SmCI with CPs, trypsin, elastase, and chymotrypsin were examined by enzyme kinetics analysis. Preincubation of the inhibitor with serine proteases over varying periods of time did not affect
Bifunctional Protease Inhibitor from Marine Invertebrate

its inhibitory capacity, suggesting that SmCI quickly binds to these enzymes, although CP inhibitory activity needed at least a minute for equilibrium to be reached before its action. In all cases, the concentrations of both protease and inhibitor were sufficiently low to generate concave inhibition curves (data not shown) indicative of reversibility (45, 46).

Values of $K_i$, adjusted according to the substrate-induced dissociation recorded for different amounts of each substrate (results not shown), indicated a strong inhibitory action on bovine pancreatic CPA, trypsin, and pancreatic elastase, with $K_i$ values between the $10^{-8}$ and $10^{-9}$ M range (Table 1). In contrast, no inhibition capacity for CPB was observed, and chymotrypsin was inhibited in smaller measure, with a $K_i$ value in the micromolar range ($1.83 \pm 0.92 \text{ M}$). Similar behavior was observed toward CPA and CPB from other species, including human forms of the enzymes (data not shown).

Primary Structure of SmCI—Edman degradation of purified natural SmCI indicated the following N-terminal sequence for the inhibitor: NH$_2$-ISVCDLPADRGQCTAYIPQWFF (22 residues released). MALDI-TOF MS revealed the incorporation of 18 carboxymethyl groups when the inhibitor was reduced and carboxymethylated, although these were absent when the intact molecule was treated, suggesting that SmCI contains 18 carboxymethyl groups when the inhibitor was reduced and carboxymethylated, although these were absent when the intact molecule was treated, suggesting that SmCI contains 18 cysteine residues involved in disulfide bonds. The reduced and carboxymethylated inhibitor was fragmented with trypsin, Lys-C, Asp-N, and Glu-C endoproteases. The resulting overlapping peptides (Fig. 2) analyzed by Edman degradation covered almost the entire sequence of SmCI. The whole sequence was subsequently completed and confirmed by cDNA cloning (see below).

Cloning and Sequence Analysis of SmCI—Using RT-PCR and nested PCR, we obtained a partial SmCI cDNA sequence from RNA that had been previously isolated as a template for the RT reaction. The product of the first PCR (S1(8–17) and R9 primers) appeared on agarose gel as a single band of around 700 bp. Nested PCR rendered a prominent PCR product of about 600 bp, which was then purified and subcloned to generate a pGEM-SmCI S1/S2 construct. Ten clones were checked by restriction analysis and sequenced using the SP6 and T7 promoter forward and reverse primers. Sequence analysis of the plasmids showed the presence of 495 bp corresponding to the SmCI cDNA sequence. By translating this nucleotide sequence, we were able to complete and confirm the SmCI protein sequence from residue 13 to the C terminus of the protein (Fig. 3).

The whole SmCI protein sequence derived from the nucleotide sequence includes 165 amino acids. The protein is rich in glycine (16 residues), glutamic acid (17 residues), and cysteine (18 residues); the latter participates in nine disulfide bridges. This sequence showed a theoretical molecular mass of 18.649 kDa, revealing a difference of 1054 Da with respect to the experimental feature determined by MALDI-TOF MS. Considering the presence of a potential N-glycosylation site (N/X/T) at Asn$^{13}$, deglycosylation assays were followed by MALDI-TOF MS. A difference of 1054 Da was found between the intact and deglycosylated form of the inhibitor, indicating that SmCI is an N-glycosylated protein.

Computer analysis of the SmCI sequence revealed a high homology with BPTI/Kunitz proteins. This facilitated a preliminary prediction of the disulfide bond pattern, fully conserved in the members of this family (52). PFAM analysis (49) indicated the presence of three BPTI/Kunitz domains in the SmCI sequences with PFAM $E$ values between $e^{-27}$ and $e^{-22}$ for each domain. NCBI nonredundant sequence database searches using PSI-BLAST revealed that SmCI most resembled proteins with more that one Kunitz domain in their structure such as the inter-$\alpha$-trypsin inhibitor of different species (horse, goat, and sheep) with $E$ values $= 3.3 \times e^{-19}$; human, bovine, pig, mouse and rat AMBP protein precursor (inter-$\alpha$-trypsin inhibitor light chain, bikunin) with $E$ values between $e^{-18}$ and $e^{-17}$; human, mouse, and rat tissue factor pathway inhibitor 1 precursor and tissue factor pathway inhibitor 2 precursor (TFPI1 and TFPI2, respectively), which have three Kunitz domains, with $E$ values of $e^{-15}$ and $e^{-10}$, respectively. Furthermore, similarities ($e^{-7}$ to $e^{-12}$) were also found with several domains of Kunitz proteins and with Kunitz-like serine protease inhibitors of species showing a different phylogenetic distribution, e.g. the sea anemone inhibitors ShPl-1 and ShPl-2 (e$^{-19}$) from Stichodactyla helianthus; Kunitz-type protease inhibitors AXPI-1 and AXP2 (e$^{-10}$ and e$^{-7}$, respectively) from Anthopleura xantho-
Bifunctional Protease Inhibitor from Marine Invertebrate

### TABLE 1

| Inhibitor | Trypsin | Pancreatic elastase | pCPA |
|-----------|---------|---------------------|------|
| SmCI      | 3.62 ± 1.39 × 10^{-9} | 8.85 ± 0.19 × 10^{-9} | 1.48 ± 0.06 × 10^{-8} |
| rSmCI     | 1.35 ± 0.13 × 10^{-9} | 8.90 ± 0.75 × 10^{-9} | 4.76 ± 0.45 × 10^{-8} |
| rSmCI N23A| 1.03 ± 0.54 × 10^{-10}| 1.15 ± 0.22 × 10^{-9} | 2.54 ± 0.13 × 10^{-8} |
| rSmCI D1-D2 | 1.22 ± 0.19 × 10^{-7} | 5.55 ± 1.01 × 10^{-8} | 1.88 ± 0.13 × 10^{-6} |
| rSmCI D2-D3 | 1.42 ± 0.30 × 10^{-10}| NI* | NI* |
| rSmCI D3 | 3.65 ± 0.77 × 10^{-8} | NI* | NI* |

*NI* means no inhibition.

### FIGURE 2

Protein sequence determination of SmCI. The N-terminal sequence (shaded) was obtained by directly sequencing the uncleaved inhibitor by Edman degradation. The peptide products of enzymatic hydrolysis of the reduced and carboxymethylated whole inhibitor, sequenced by Edman degradation. The peptide products of enzymatic hydrolysis (shown in Fig. 3). The N-terminal sequence of protein is displayed in parentheses. The rest of the sequence is underlined. The consensus polyadenylation signal is underlined. The terminal codon is marked by an asterisk.

The nucleotide sequence of SmCI was deposited in GenBank™ under accession number AM283480.

### FIGURE 3

Nucleotide and deduced amino acid sequence of SmCI. Residues obtained by DNA translation (from residue 13) are shown in bold; the amino acids were determined by direct protein sequencing. The peptide products of enzymatic hydrolysis (shown in Fig. 3).

### FIGURE 4

Structure in the first two domains and feature an eight-stranded sandwich sheet toward the C terminus, which is typical of BPTI/Kunitz protein sequences. The central anti-parallel α-helix in the second domain is important for CPs. The third domain bears the typical reactive site for trypsin inhibition with a lysine residue at P1 and enzyme recognition sites around the P1-P1' reactive site are depicted for the three domains in this figure.

### Modeling the Three-dimensional Structure of SmCI Domains

To account for the high degree of similarity between SmCI and BPTI/Kunitz protein sequences along with the absence of three-dimensional structures established for proteins with three Kunitz domains in their structure, we computationally modeled the individual SmCI domains. The three-dimensional crystal structures used as templates were selected on the basis of their three-dimensional JURY scores (above 45%) and high resolution (below 2.5 Å) (53). To construct specific models for SmCI-D1, SmCI-D2, and SmCI-D3, we used as template the structure of the second domain of human TFPI1 (Protein Data Bank code 1TFX), the second domain of human bikunin (Protein Data Bank code 1BIK), and the first domain of human TFPI2 (Protein Data Bank code 1ZRO), respectively.

Based on these templates, we generated 100 models for each domain using the MODELLER program (54) and quality packing values below −5.0. Fig. 5 shows the best calculated models for each domain with average values for the quadratic medium deviation of 100 models equal to 0.223 ± 0.04, 0.199 ± 0.031, and 0.252 ± 0.024 for SmCI-D1, SmCI-D2, and SmCI-D3, respectively. Each domain is folded in a central anti-parallel β-sheet with one α-helix toward the C terminus, which is typical of BPTI/Kunitz proteins. However, according to the models, the domains differ in their N terminus, which lack a regular structure in the first two domains and feature an α-helix in the third domain, instead of the 310 helix commonly found in this family of proteins.

### FIGURE 5

Comparison of the three-dimensional structures of SmCI-D1, SmCI-D2, and SmCI-D3 with that of the second domain of human TFPI1 (Protein Data Bank code 1ZRO). The secondary structure of SmCI-D1, SmCI-D2, and SmCI-D3 is shown as a ribbon diagram. The secondary structure of SmCI-D1, SmCI-D2, and SmCI-D3 is shown as a ribbon diagram.
FIGURE 4. SmCI primary structure and domain arrangement. A, schematic basic structure of BPTI/Kunitz family domains; the consensus sequence of this family identified in each domain is indicated by the dotted box. The P3-P3 residues of each domain’s reactive site appear in the solid line box. P1-P1 residues are shown in boldface type. B, amino acid sequences of SmCI domains; the consensus sequence of this family identified in each domain is indicated by the dotted box. The P3-P3 residues of each domain’s reactive site appear in the solid line box. P1-P1 residues are shown in boldface type. C, diagram showing SmCI’s primary structure composed of three Kunitz domains. The nine disulfide bridges (three for each domain) are designated according to the BPTI/Kunitz disulfide pattern (Cys1-Cys15, Cys9-Cys16, and Cys4-Cys9). The glycosylated Asn is marked with an asterisk. Arrows point to the P1-P1 site of each domain. The canonical loops (P3-P3) of each Kunitz domain are indicated by a curve. This sequence was deposited in the UNIPROT Data bank under the access number P84875 and identifier PCPI_SABMA.

FIGURE 5. Three-dimensional structure models of the individual SmCI domains. Each domain shares the typical three-dimensional structure of BPTI/Kunitz proteins as follows: a central anti-parallel β-sheet and an α-helix toward the C terminus. Note that the three domains differ most in terms of their N terminus, which lacks any regular structure in the first and second domains and shows an α-helix in the third domain. The P1 residues at the active site for each domain are shown as sticks. In all cases, reactive loops maintain the canonical conformation. Images were prepared using the PyMOL version 0.99 program.
Bifunctional Protease Inhibitor from Marine Invertebrate

(a) Glycosylation

(b) N-terminal

(c) C-terminal

FIGURE 6. Comparing the C-terminal sequences of metallo-CP inhibitors. Aligned inhibitors, SmCI (PCPI_SABMA), carboxypeptidase inhibitor from S. magnifica; TCI (TCI1_RHIBU), carboxypeptidase inhibitor from A. bursa (9); HITCI (AB836_HAELO), carboxypeptidase inhibitor from H. longicornis (10); MPCI (MCPI_SOLLC) carboxypeptidase inhibitor from Solanum lycopersicum (6); PCI (MCPI_SOLTU) carboxypeptidase inhibitor from Solanum tuberosum (3, 4); LCI (MCPI_HIRM) carboxypeptidase inhibitor from H. medicinals (8); A. xanthogrammica

TABLE 2

| Inhibitors       | P3     | P2     | P1     | P1'    | P2'    | P3'    |
|------------------|--------|--------|--------|--------|--------|--------|
| SmCI D1          | Glu    | Cys    | Thr    | Ala    | Tyr    | Ile    |
| SmCI D2          | Pro    | Cys    | Arg    | Val    | Ser    | Ala    |
| SmCI D3          | Pro    | Cys    | Lys    | Gly    | Ser    | Phe    |
| AXPI1_ANTAF      | Pro    | Cys    | Arg    | Ala    | Ala    | Val    |
| AXPI2_ANTAF      | Val    | Cys    | Arg    | Ala    | Ala    | Val    |
| IPS2_ANESU       | Pro    | Cys    | Arg    | Ala    | Arg    | Phe    |
| ISH1_STOHE       | Arg    | Cys    | Lys    | Gly    | Tyr    | Phe    |
| ISH2_STOHE       | Arg    | Cys    | Lys    | Gly    | Tyr    | Phe    |
| AceKI-1          | Glu    | Cys    | Met    | Ala    | Phe    | Phe    |
| Q8MrR6_HAEIR     | Pro    | Cys    | Arg    | Lys    | Ser    | Asp    |
| BMTIA_BOOMI-D1   | Pro    | Cys    | Arg    | Ala    | Ala    | Leu    |
| BMTIA_BOOMI-D2   | Pro    | Cys    | Leu    | Ala    | Ala    | Met    |
| CSTI_BOMMO       | Pro    | Cys    | Met    | Gly    | Thr    | Phe    |
| AMBP_HUMAN-D1    | Pro    | Cys    | Arg    | Ala    | Phe    | Phe    |
| AMBP_BOVIN-D1    | Pro    | Cys    | Leu    | Gly    | Leu    | Phe    |
| SPIT2_HUMAN-D1   | Arg    | Cys    | Arg    | Ala    | Ser    | Met    |
| TFPPI2_HUMAN-D1  | Pro    | Cys    | Arg    | Ala    | Leu    | Leu    |
| BPTI_BOVIN       | Pro    | Cys    | Lys    | Ala    | Arg    | Ile    |

a glycine at P1', consistent with the inhibitory activity toward trypsin observed for this domain.

Heterologous Expression and Characterization of Recombinant Inhibitors—Using the cloning strategies described under “Experimental Procedures,” we produced the different recombinant forms of the inhibitor. By combined rSmCI overexpression and purification, we produced a fraction that was highly enriched in the active recombinant protein showing several signals around 20 kDa in MALDI-TOF MS spectra (results not shown). These signals probably corresponded to different glycosylation grades of the protein, frequently described in the P. pastoris system (65). Automatic Edman degradation analysis revealed a single N-terminal sequence (ISVCDLPA DR) matching the natural SmCI N-terminal sequence.

In contrast, rSmCI N23A, made to avoid the glycosylation, showed a single peak with a molecular mass of 18,606 Da by MALDI-TOF MS, which correspond to the theoretical molecular mass for this alanine mutant determined by ExPaSy server. The asparagine residue mutated was confirmed by Edman sequencing.

Purified rSmCI and rSmCI N23A revealed the same inhibitory capacity and specificity toward proteases as the natural inhibitor. Dose-effect assays of the inhibitors against metallo-carboxypeptidase A showed a concave kinetic behavior curve for this enzyme (Fig. 7A). The recalculated $K_i$ values (considering substrate dissociation) against CPA, trypsin, and pancreatic elastase are summarized in Table 1. The recombinant tri-domain (glycosylated and nonglycosylated) and natural forms showed similar $K_i$ values against pancreatic CPA and serine proteases confirming their bifunctional properties similar to the natural SmCI inhibitor.

When the bi-domain proteins rSmCI D1-D2 and rSmCI D2-D3 were overexpressed and purified, we got two proteins with molecular masses of 11,949 and 12,549 Da, respectively, and N-terminal sequences that correspond with both proteins. rSmCI D1-D2 was able to inhibit trypsin and pancreatic elastase with $K_i$ values of $1.20 \times 10^{-7}$ and $5.55 \times 10^{-10}$ M, respectively, and for CPA, rSmCI D1-D2 showed $K_i = 1.96 \times 10^{-6}$ M, indicating less strength of inhibition against the three enzymes than that obtained for tri-domain inhibitor forms (Fig. 7B and Table 1). rSmCI D2-D3 and rSmCI D3 only inhibited trypsin with $K_i$ values of $1.4 \times 10^{-9}$ and $3.05 \times 10^{-8}$ M, respectively, displaying null inhibitory activity on other serine proteases or on CPs (Table 1).

DISCUSSION

In this study, we identified and characterized (at both the protein and DNA levels) a novel metallo-carboxypeptidase inhibitor (SmCI) isolated from the tentacle crown of the marine annelid S. magnifica. By using several complementary strategies, we were able to establish its complete SmCI protein sequence. In addition, an N-glycosylation site at residue Asn-23 was predicted and confirmed experimentally in the natural form by deglycosylation assays.

Functionally, SmCI is a tight-binding inhibitor that inhibits both carboxypeptidase A and serine proteases, such as trypsin and pancreatic elastase, with equilibrium dissociation constants in the nanomolar range, whereas chymotrypsin is inhibited with a $K_i$ value only in the micromolar range. These functional characteristics were reproduced by the recombinant tri-domain forms (glycosylated and nonglycosylated), confirming that they are correctly folded and functionally equivalent to the protein isolated from the tentacle crown. Nevertheless, the results obtained in this work suggest that in case of SmCI N23A, a nonglycosylated form of the protein by elimination of the sugar moiety, it displays an increment of its trypsin inhibitory capability, probably reflecting a better enzyme-inhibitor fit.
However, this amino acid residue, which bears a threonine, a residue not frequently found at this position. The role of threonine at position P1 in the BPTI/Kunitz protease inhibitor family is exemplified by bovine pancreatic trypsin inhibitor, has been the subject of intense investigations. Numerous variants of these protein inhibitors have been isolated from different tissues and animal sources, spanning from turtle egg white and snake venom to the major organs of ruminant mammals. The specificity of each of these protease inhibitors is mainly dependent on the structure of the members of this family, including the same spacing of the six conserved cysteine residues, as well as the consensus sequence FXXGCXGNXXN around the fourth cysteine residue. The three-dimensional structures generated for each domain by comparative modeling also displayed the typical structural features of BPTI/Kunitz inhibitors, as the most representative canonical protease inhibitors.

The BPTI/Kunitz protease inhibitor family, exemplified by bovine pancreatic trypsin inhibitor, has been the subject of intense investigations. Numerous variants of these protein inhibitors have been isolated from different tissues and animal sources, spanning from turtle egg white and snake venom to the major organs of ruminant mammals. The specificity of each of these protease inhibitors is mainly dependent on the nature of the residue at position P1 of the active site. Effectively, two of SmCI domains (the second and third) have basic P1 residues, such as arginine and lysine, respectively, as do most BPTI/Kunitz trypsin inhibitors; although the first SmCI domain bears a threonine, a residue not frequently found at this position.

The role of threonine at position P1 in the BPTI/Kunitz protease inhibitor family has not yet been well defined. However, this amino acid residue, which bears a β-branched side chain, has been described to bind to the S1 pocket of trypsin in a particularly weak mode. In contrast, studies involving the use of peptide combinatorial libraries have suggested that threonine could promote the inhibition of elastases. It therefore seems reasonable to suggest that the SmCI-D1 (domain 1, containing such threonine) could contribute to a lesser extent to inhibiting trypsin and be responsible for elastase inhibition. The additional presence of a hydrophobic residue at P2', able to interact with the S2' subsite of the enzyme, could favor elastase inhibition by this first SmCI domain. The inhibition of pancreatic elastase by the rSmCI D1-D2 and its lack by the rSmCI D2-D3 indicates that the D1 domain is probably responsible for the inhibition of this enzyme, also suggesting an important role of the P1 threonine residue in this event. The important role of P1 threonine residue for elastase inhibition has been demonstrated by the BPTI mutants and is in agreement with our finding. The BPTI K15T mutant (with threonine at P1 position) showed a $K_i$ value 3 orders lower than natural BPTI.

A remarkable characteristic of the second SmCI domain is the presence of valine at position P1'. Indeed, almost all Kunitz domains feature an alanine or glycine residue at this position, although other small side chain amino acid residues such as aspartate and serine have often been found at this position. These residues were found to be present in β-branched side chain residues such as valine causes a decrease in the energy of interaction with trypsin (also described for chymotrypsin), destabilizing the complex. Hence, the presence of lysine and glycine at P1 and P1', respectively, of the third domain indicates a likely greater contribution to trypsin inhibition by this domain, a fact confirmed experimentally with the lower $K_i$ value determined for rSmCI D2-D3 in comparison with that obtained for rSmCI D1-D2.

BPTI/Kunitz domains, along with those of other inhibitors of the canonical class (Kazal, Bowman-Birk, grasshopper, etc.), can be found as repeats forming a multidomain single chain inhibitor, able to independently interact with several proteases at their reactive sites. From this perspective, it is obvious that SmCI can be classified as a multidomain inhibitor of serine proteases with three reactive sites (one at each domain), each one probably being responsible for the preferential inhibition of different serine proteases.
Bifunctional Protease Inhibitor from Marine Invertebrate

...
peptidase inhibitors from *Ascaris suum*. The primary structure. Arch. Biochem. Biophys. 270, 153–161.
7. Sanglas, L., Aviles, F. X., Huber, R., Gomis-Ruth, F. X., and Arolas, J. L. (2009) Mammalian metalloproteinase inhibition at the defense barrier of *Ascaris* parasite. Proc. Natl. Acad. Sci. U.S.A. 106, 1743–1747.
8. Reverter, D., Vendrell, J., Canals, F., Horstmann, J., Avilés, F. X., Fritz, H., and Sommerhoff, C. P. (1998) A carboxypeptidase inhibitor from the medical leech *Hirudo medicinalis*. Isolation, sequence analysis, cDNA cloning, recombinant expression, and characterization. J. Biol. Chem. 273, 32927–32933.
9. Arolas, J. L., Lorenzo, J., Rovira, A., Castellà, J., Avilés, F. X., and Sommerhoff, C. P. (2005) A carboxypeptidase inhibitor from the tick *Rhipicephalus bursa*. Isolation, cDNA cloning, recombinant expression, and characterization. J. Biol. Chem. 280, 3441–3448.
10. Gong, H., Zhou, J., Liao, M., Hatta, T., Harnnoi, T., Umemiya, R., Inoue, N., Yuan, X., and Fujisaki, K. (2007) Characterization of a carboxypeptidase inhibitor from the tick *Haemaphysalis longicornis*. J. Insect Physiol. 53, 1079–1087.
11. Normant, E., Martres, M. P., Schwartz, J. C., and Gros, C. (1995) Purification, cDNA cloning, functional expression, and characterization of a 26-kDa endogenous mammalian carboxypeptidase inhibitor. Proc. Natl. Acad. Sci. U.S.A. 92, 12225–12229.
12. Liu, Q., Yu, L., Gao, J., Fu, Q., Zhang, J., Zhang, P., Chen, J., Zhao, S. (2000) Cloning, tissue expression pattern, and genomic organization of latexin, a human homologue of rat carboxypeptidase A inhibitor. Mol. Biol. Rep. 27, 241–246.
13. Arolas, J. L., Popowicz, G. M., Schwartz, J. C., and Gros, C. (1995) Purification, cDNA cloning, functional expression, and characterization of a 26-kDa endogenous mammalian carboxypeptidase inhibitor. Proc. Natl. Acad. Sci. U.S.A. 102, 3978–3983.
14. Guasch, A., Coll, M., Avilés, F. X., and Huber, R. (1992) Three-dimensional structure of porcine pancreatic carboxypeptidase A. A comparison of the A and B zymogens and their determinants for inhibition and activation. J. Mol. Biol. 214, 141–157.
15. Ventura, S., Villegas, V., Sterner, J., Larson, J., Vendrell, J., Hershberger, C. L., and Avilés, F. X. (1999) Mapping the pro-region of carboxypeptidase B by protein engineering. Cloning, overexpression, and mutagenesis of the procarboxypeptidase A4 with its endogenous protein inhibitor, latexin. Proc. Natl. Acad. Sci. U.S.A. 102, 14982–14987.
16. Delfín, J., González, E., and Chávez, M. (1994) Proteinase inhibitor from *Stichodactyla helianthus*. Purification, characterization, and immobilization. Arch. Med. Res. 25, 199–204.
17..Lenarcic, B., Ritonja, A., Strakelj, B., Turk, B., and Turk, V. (1997) Equistatin, a new inhibitor of cysteine proteinases from *Actinia equina*, is structurally related to thyroglobulin-type-1 domain. J. Biol. Chem. 272, 13899–13903.
18. Strukelj, B., Lenarcic, B., Gruden, K., Pungercar, J., Rogelj, B., Turk, V., Bosch, D., and Longstra, M. A. (2000) Equistatin, a protease inhibitor from the sea anemone *Actinia equina*, is composed of three structural and functional domains. Biochem. Biophys. Res. Commun. 269, 732–736.
19. Mebs, D., and Gebauer, E. (1980) Isolation of proteinase inhibitory, toxic, and hemolytic polypeptides from a sea anemone. Stoichactis sp. Toxicon 18, 97–106.
20. Chávez, M. A., Gil, Sh., Fernández, A., Huerta, V., Pascual, I., Abreu, L., Morera, V., Saroyán, A., Delfín, J., Padrón, G., Cisneros, M., Joseph, P., Charli, J. L., and Díaz, I. (1998) Purification and partial characterization of a proteinase inhibitor from the sea anemone *Condylactis gigantea*. Toxicon 36, 1275.
21. Fritz, H., Brey, B., and Béress, L. (1972) Polyvalent isoinhibitors for trypsin, chymotrypsin, plasmin, and kallikreins of sea anemones (*Anemone sulcata*), isolation, inhibitory behavior, and amino acid composition. Hoppe-Seyer’s Z. Physiol. Chem. 353, 19–30.
22. Kolkenbrock, H., and Tschesche, H. (1987) A new inhibitor of elastase from the sea anemone (*Anemone sulcata*). Biochem. Biol. Hoppe-Seyer *368*, 93–99.
Bifunctional Protease Inhibitor from Marine Invertebrate

chymosins A and B on a synthetic chromophoric hexapeptide. Biochim. Biophys. Acta 612, 410–420
44. Chase, T., Jr., and Shaw, E. (1967) p-Nitrophenyl-p'-guanidinobenzoate HCl. A new active site titrant for trypsin. Biochem. Biophys. Res. Commun. 29, 508–514
45. Bieth, J. G. (1995) Theoretical and practical aspects of proteinase inhibition kinetics. Methods Enzymol. 248, 59–84
46. Williams, J. W., and Morrison, J. F. (1979) The kinetics of reversible tight-binding inhibition. Methods Enzymol. 63, 437–467
47. Bergmeyer, H. U., Gawehn, K., and Grassl, M. (1974) in Methods of Enzymatic Analysis (Bergmeyer, H. U., ed) Vol. 1, 2nd Ed., pp. 436–437, Academic Press, Inc., New York
48. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402
49. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., and Sonnhammer, E. L. (2002) The Pfam protein families database. Nucleic Acids Res. 30, 276–280
50. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882
51. Karlin, S., and Altschul, S. F. (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc. Nat. Acad. Sci. U.S.A. 87, 2264–2268
52. Hulo, N., Bairoch, A., Bilouli, V., Cerutti, L., De Castro, E., Landgiedжи-Genaveux, P. S., Pagni, M., and Sigrist, C. J. (2006) The PROSITE database. Nucleic Acids Res. 34, D227–D230
53. Kaján, L., and Rychlewski, L. (2007) Evaluation of three-dimensional Jury on CASP7 models. BMC Bioinformatics 8, 304
54. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., and Karplus, M. (1995) Evaluation of comparative protein modeling by MODELLER. Proteins 23, 318–326
55. Huerta, V., Morera, V., López, N., Betancourt, L., Besada, V., Padrón, G., Lima, G., Chávez, M. A., Delfin, J., and Diaz, J. (1988) Characterization and three-dimensional model of a new proteinase inhibitor isolated from Stichodactyla helianthus. Biotechnologia Aplicada 15, 108–110
56. Millstone, A. M., Harrison, L. M., Bungiro, R. D., Kuzmic, P., and Cappello, M. (2000) A broad spectrum Kunitz-type serine proteinase inhibitor secreted by the hookworm Anicylostoma ceylanicum. J. Biol. Chem. 275, 29391–29399
57. Sant’Anna Azolli, S., Sasaki, S. D., Torquato, R. I., Andreotti, R., Andreotti, E., and Tanaka, A. S. (2003) Rhizophaeis sanguineus trypsin inhibitors present in the tick larvae. Isolation, characterization, and partial primary structure determination. Arch. Biochem. Biophys. 417, 176–182
58. Tanaka, A. S., Andreotti, R., Gomes, A., Torquato, R. J., Sampaio, M. U., and Sampaio, C. A. (1999) A double-headed serine proteinase inhibitor, human plasma kallikrein and elastase inhibitor, from Boophilus microplus larvae. Immunopharmacology 45, 171–177
59. Fritz, H., and Wunderer, G. (1983) Biochemistry and application of aprotinin, the kallikrein inhibitor from bovine organs. Arzneimittelforschung 33, 479–494
60. Kurioka, A., Yamazaki, M., and Hirano, H. (1999) Primary structure and possible functions of a trypsin inhibitor of Bombbyx mori. Eur. J. Biochem. 259, 120–126
61. Xu, Y., Carr, P. D., Guss, J. M., and Ollis, D. L. (1998) The crystal structure of bikunin from the inter-α-trypsin inhibitor. A serine proteinase inhibitor with two Kunitz domains. J. Mol. Biol. 276, 955–966
62. Hochstrasser, K., Albrecht, G., Schönberger, O. L., and Wachter, E. (1983) Kunitz-type proteinase inhibitors derived by limited proteolysis of the inter-α-trypsin inhibitor, VII. Characterization of the bovine inhibitor as double-headed trypsin-elastase inhibitor. Hoppe-Seyler’s Z. Physiol. Chem. 364, 1689–1696
63. Marlor, C. W., Delaria, K. A., Davis, G., Muller, D. K., Gree, J. M., and Tamburini, P. P. (1997) Identification and cloning of human placental bikunin, a novel serine proteinase inhibitor containing two Kunitz domains. J. Biol. Chem. 272, 12202–12208
64. Chand, H. S., Schmidt, A. E., Bajaj, S. P., and Kisiel, W. (2004) Structure-function analysis of the reactive site in the first Kunitz-type domain of human tissue factor pathway inhibitor-2. J. Biol. Chem. 279, 17500–17507
65. Cereghino, J. L., and Cregg, J. M. (2000) Heterologous protein expression in the methylotrophic yeast Pichia pastoris. FEMS Microbiol. Rev. 24, 45–66
66. Bode, W., and Huber, R. (1992) Natural protein proteinase inhibitors and their interaction with proteinases. Eur. J. Biochem. 204, 433–451
67. Laskowski, M., Jr., Qasmim, M. A., and Lu, S. M. (2000) in Protein-Protein Recognition (Kleanthous, C., ed) pp. 228–279, Oxford University Press, Oxford
68. Czapinska, H., Otlewski, J., Krzywda, S., Sheldrick, G. M., and Jaskólski, M. (2000) High resolution structure of bovine pancreatic trypsin inhibitor with altered binding loop sequence. J. Mol. Biol. 295, 1237–1249
69. Laskowski, M., Jr., and Kato, I. (1980) Protein inhibitors of proteinases. Annu. Rev. Biochem. 49, 593–626
70. Grzesiak, A., Helland, R., Smals, A. O., Krowarsch, D., Dadlez, M., and Otlewski, J. (2000) Substitutions at the P(1) position in BPTI strongly affect the association energy with serine proteinases. J. Mol. Biol. 301, 205–217
71. Otlewski, J., Krowarsch, D., and Apostoluk, W. (1999) Protein inhibitors of serine proteinases. Acta Biochim. Pol. 46, 531–565
72. Krowarsch, D., Dadlez, M., Buczek, O., Krokoszynska, I., Smalas, A. O., and Otlewski, J. (1999) Interscaffolding additivity: binding of P1 variants of bovine pancreatic trypsin inhibitor to four serine proteinases. J. Mol. Biol. 289, 175–186
73. Shiù, S., Stamos, J., Kirchhofer, D., Fan, B., Wu, J., Corpuz, R. T., Santell, L., Lazarus, R. A., and Eigenbrot, C. (2005) Conformational lability in serine proteinase active sites. Structures of hepatocyte growth factor activator (HGFα) alone and with the inhibitory domain from HGFα-inhibitor-1B. J. Mol. Biol. 346, 1335–1349
74. McBride, J. D., Freeman, H. N., and Leatherbarrow, R. J. (1999) Selection of human elastase inhibitors from a conformationally constrained peptide library. Eur. J. Biochem. 266, 403–412
75. Hiraga, K., Suzuki, T., and Oda, K. (2000) A novel double-headed proteinaceous inhibitor for metalloproteinase and serine proteinase. J. Biol. Chem. 275, 25173–25179
76. Franco, O. L., Grossi de Sá, M. F., Sales, M. P., Mello, L. V., Oliveira, A. S., Rigden, D. J. (2002) Overlapping binding sites for trypsin and papain on a Kunitz-type proteinase inhibitor from Prosopis juliflora. Proteins 49, 335–341
77. Zemke, K. J., Müller-Fahrnow, A., Jany, K. D., Pal, G. P., and Saenger, W. (1991) The three-dimensional structure of the bifunctional proteinase K1a-amylase inhibitor from wheat (PK13) at 2.5 Å resolution. FEBS Lett. 279, 240–242
78. Vallée, F., Kadziola, A., Bourne, Y., Juy, M., Rodenburg, K. W., Svensson, B., and Haser, R. (1998) Barley α-amylase bound to its endogenous proteinase inhibitor BASI. Crystal structure of the complex at 1.9 Å resolution. Structure 6, 649–659
79. Arolas, J. L., Lorenzo, J., Rovira, A., Vendrell, J., Aviles, F. X., and Ventura, S. (2004) Secondary binding site of the potato carboxypeptidase inhibitor. Biochemistry 43, 7973–7982