Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 67 (2006) 362–370
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Low molecular weight squash trypsin inhibitors from *Sechium edule* seeds

Hélen J. Laure [a,b], Vítor M. Faça [a,b], Clarice Izumi [b,c], Júlio C. Padovan [b,d], Lewis J. Greene [b,c,*]

[a] *Departamento de Bioquímica, Área de Biologia Molecular, Universidade Federal de São Paulo, Escola Paulista de Medicina, São Paulo, SP, Brazil*
[b] *Centro de Química de Proteínas, Faculdade de Medicina de Ribeirão Preto, Brazil*
[c] *Departamento de Biologia Celular, Molecular e Bioagentes Patogênicos, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes, 3900, 14049-900 Ribeirão Preto, SP, Brazil*
[d] *Laboratory for Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, NY 10021-6399, USA*

Received 11 February 2005; received in revised form 10 September 2005
Available online 9 January 2006

Abstract

Nine chromatographic components containing trypsin inhibitor activity were isolated from *Sechium edule* seeds by acetone fractionation, gel filtration, affinity chromatography and RP-HPLC in an overall yield of 46% of activity and 0.05% of protein. The components obtained with highest yield of total activity and highest specific activity were sequenced by Edman degradation and their molecular masses determined by mass spectrometry. The inhibitors contained 31, 32 and 27 residues per molecule and their sequences were: SETI–IIa, EDRKCPKILMRCKRDSDCLAKCTCQESGYCG; SETI–IIb, EEDRKCPKILMRCKRDSDCLAKCTCQESGYCG and SETI-V, CPRILMKCKLDTDCFPTCTCRPSGFCG. SETI–IIa and SETI–IIb, which differed by an amino-terminal E in the IIb form, were not separable under the conditions employed. The sequences are consistent with consensus sequences obtained from 37 other inhibitors: CPriI1meCk_DSDCla_C_C_G_CG, where capital letters are invariant amino acid residues and lower case letters are the most preserved in this position. SETI–II and SETI–V form complexes with trypsin with a 1:1 stoichiometry and have dissociation constants of 5.4×10^{-11} M and 1.1×10^{-9} M, respectively.

Keywords: *Sechium edule* ("Chuchu"); Cucurbitaceae; Serine protease; Squash trypsin inhibitor; Amino acid sequence; Trypsin activity; Mass spectrometry

1. Introduction

Protein protease inhibitors are divided into four superfamilies, i.e., serine proteases, metalloproteases, cysteine proteases, and aspartic proteases (Laskowski and Kato, 1980; Bode and Huber, 1993). They contain 29–190 amino acid residues/molecule and do not represent a uniform evolutionary group. Inhibitors have a compact shape, a hydrophobic nucleus, and some have several disulfide bridges/molecule (Bode and Huber, 1992, 1993). The plant serine protease inhibitors are divided into families which include Kunitz, Bowman-Birk, Potato I and II and Squash, among others (Otlewski, 1993).

Squash-type trypsin inhibitors are the smallest serine protease inhibitors known. They have been purified from members of the Cucurbitaceae family, such as *Cucurbita*, *Cucumis* and *Momordica*. These inhibitors form 1:1 complexes with trypsin, Factor XII of plasma and, in some cases, with elastase. The high association constants with bovine trypsin (10^8–10^{11} M^{-1}) identify them as strong protease inhibitors (Wieczorek et al., 1985; Krishnamoorthi et al., 1990). Their polypeptide chains are usually comprised of 27–34 amino acid residues with, a high relative

* Corresponding author. Tel./fax: +55 16 2101 9366.
E-mail address: ljgreene@fmrp.usp.br (L.J. Greene).

0031-9422/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2005.11.016

content of cysteinyl residues, six in the case of squash inhibitors. The relative positions of the cysteines, as well as their relative pairing, are identical among the other members of the squash trypsin inhibitor family.

The reactive site is close to the amino terminus and contains either an arginyl (Arg) or a lysyl (Lys) residue followed by a leucine (Leu). It is part of a loop between two cysteinyl (Cys) residues with the ε-amino or guanidino-group of the reactive site oriented outward from the body of the inhibitor in order to interact with the active site of trypsin (Hider et al., 1987).

Squash trypsin inhibitors have been synthesized and their disulfide bonds oxidized to provide an active trypsin inhibitor (Le-Nguyen et al., 1989). The P1 residue of serine protease inhibitors may be substituted without loss of inhibitory activity (Laskowski and Kato, 1980), and synthesis of squash inhibitors with appropriate amino acid substitutions in P1 can modify specificity (Rolka et al., 1989, 1991). This is important because protease inhibitors are used as drugs for treatment of some diseases such as AIDS (McQuade et al., 1990).

These features of squash inhibitors stimulated us to isolate and characterize trypsin inhibitors from *Sechium edule* (a variety of Cucurbitaceae indigenous to Brazil called "chuchu"), in order to use them in a synthesis program for the development of new protease inhibitors of biological interest by modifying their amino acid sequences.

2. Results and discussion

2.1. Preparation of inhibitors

The extract of *S. edule* seeds (3.2 kg) with 100 mM ammonium acetate contained 3525 IU (units of trypsin inhibitor activity) and 35.2 g protein. Acetone (Me_2CO) fractionation (60–90%, v/v) gave a yellow liquid phase, immiscible with the H_2O–Me_2CO phase (60–90%, v/v), which contained 4.7 g protein (13.3%) and 73% of the trypsin inhibitor activity. The activity was eluted from a Sephadex G-25 column from 600 to 800 ml after albumin (520–660 ml), with a recovery of 84% and 2.1% (752 mg) of protein.

After the G-25 column chromatography, the active fraction was applied to a Trypsin-Sepharose affinity column, with recoveries of 61% and 0.25% (89 mg) for inhibitor activity and protein, respectively. A final purification step on a C_{18} column separated the mixture into at least 11 components, with 9 having inhibitor activity (Fig. 1). Peaks II and V were purified 844 and 957 times on the basis of specific trypsin inhibitor activity. The amount of protein associated with activity was 16 mg or 0.05% of the original protein in the seed extract, and contained 46% of the inhibitor activity. The trypsin inhibitors from *S. edule* were named using the general prefix SETI for *Sechium Edule Trypsin Inhibitor* followed by Roman numerals according to Wieczorek et al. (1985).

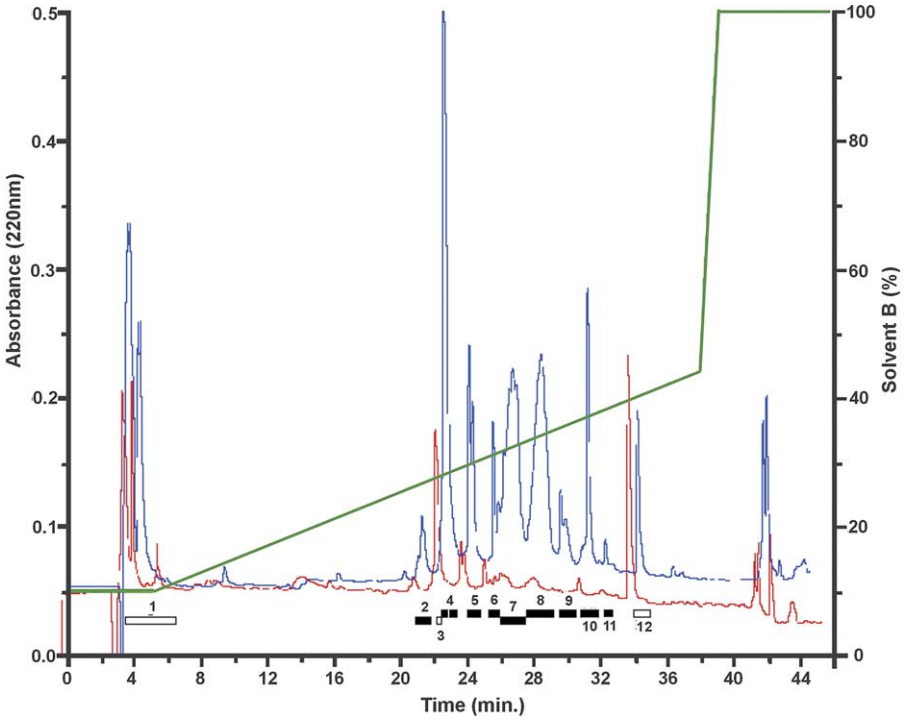

Fig. 1. Purification of *S. edule* trypsin inhibitors by RP-HPLC after affinity chromatography. Separation was performed on a Vydac C_{18} column (4.6 × 150 mm, 10 mm particle size) in 0.1% trifluoroacetic acid with an increasing linear acetonitrile (MeCN) gradient at a flow rate of 1.0 ml/min. The eluate was monitored simultaneously at 220 and 280 nm. The black rectangles indicate fractions containing trypsin inhibitor activity.

2.2. SETI–II

The amino acid composition of SETI–II reported in Table 1 corresponds to a minimum of 31 amino acid residues. The presence of 6 cysteins per mole of inhibitor was demonstrated by Edman degradation of the reduced and pyridylethylated inhibitor (PE-inhibitor), in which PTH-PEC was eluted before PTH-Pro from the HPLC column.

The sequence obtained for unmodified SETI–II (data not shown) was the same as that of PE-SETI–II (Table 2). There was a lag in the release of PTH-amino acids, which started at the second cycle and continued to the end of the sequence. We hypothesized that this lag was due to a second inhibitor that co-eluted with SETI–II, which contained an extra residue of glutamic acid at its N-terminus. This would account for the results shown in Table 2.

MALDI-TOF/MS analysis of unmodified SETI–II and PE-SETI–II showed that two species were present in each sample. The average molecular mass of 3535.1 u for unmodified SETI–II is in good agreement with the average mass of 3535.1 u obtained from the sequence data, with the assumption that all six cysteinyl residues are in disulfide bonds. The second species present had an average molecular mass of 3664.3 u, showing a mass difference of 129.2 ± 0.1 u, the value expected for an extra dehydrated glutaminyl residue of 129.1 u (Fig. 2, insets in panels A and B). For the pyridylethylated preparation, the two peaks present had average molecular masses of 4172.1 u and 4301.2 u, with a mass difference of 129.1 ± 0.1 u.

Using electrospray ionization MS with an ion trap enabled us to obtain isotopic resolution of the quadruply charged ions of the two species, now denoted SETI–IIa and SETI–IIb, with the latter bearing the higher mass. The mass differences observed in the monoisotopic mass measurements were 128.93 ± 0.08 u in both preparations (unmodified and pyridylethylated), which agree with the monoisotopic mass of 129.04 u for a dehydrated glutaminyl residue (Fig. 2).

Additional evidence for the presence of two species for SETI–II was obtained by measurement of the cyanogen bromide products. Cleavage at the single methionyl residue at position 8 generated one peptide with an average mass of 2852.5 u which mapped to the C-terminal region of SETI–II with an experimental error of $\Delta M = +0.1$ u, and two peptides with average masses of 1289 u and 1418.7 u. These peptides can be assigned to the N-terminal of SETI–II with mass errors of −0.1 u for SETI–IIa

Table 1
Amino acid composition of *S. edule* trypsin inhibitors

AA	Amino acid composition			
	SETI–II (unmodified)	PE-SETI–II	SETI–V (unmodified)	PE-SETI–V
Asp	2.97 (3)	2.83 (3)	2.22 (2)	2.03 (2)
Glu	3.13 (3)	3.45 (3)	0.65 (0)	0.49 (0)
Ser	1.72 (2)	1.79 (2)	1.05 (1)	1.05 (1)
Gly	2.13 (2)	1.97 (2)	2.35 (2)	2.27 (2)
His	0.20 (0)	nd (0)	0.13 (0)	nd (0)
Arg	2.88 (3)	2.63 (3)	2.14 (2)	1.18 (1)
Thr	1.07 (1)	0.97 (1)	2.59 (3)	2.63 (3)
Ala	1.21 (1)	0.96 (1)	0.22 (0)	0.14 (0)
Pro	1.00 (1)	1.00 (1)	2.50 (3)	2.08 (2)
Tyr	0.74 (1)	0.77 (1)	nd (0)	nd (0)
Met	0.74 (1)	0.74 (1)	0.51 (1)	1.04 (1)
Cys	3.47 (6)	7.51 (6)	2.33 (6)	6.26 (5)
Ile	0.81 (1)	0.92 (1)	0.90 (1)	0.89 (1)
Leu	1.84 (2)	1.83 (2)	1.58 (2)	1.91 (2)
Phe	nd (0)	nd (0)	1.58 (2)	1.80 (2)
Lys	4.00 (4)	3.60 (4)	2.08 (2)	1.94 (2)
Sum	31	31	27	24

The amino acid composition of each inhibitor was determined before and after reduction and pyridylethylation (PE).
The number of Cys residues was determined on the basis of sequence data.
Residues with molar ratios <0.10 were not included in the table.
Numbers in parenthesis were confirmed by sequencing.
nd = not determined.

Table 2
Sequence of pyridylethylated SETI–II (PE-SETI–II) and unmodified SETI–V

Sample cycle	PE-SETI–IIa	PE-SETI–IIb	SETI–V	
1	E	31.6	I 422.9	C (−)
2	D 19.9	E 15.1	L 306.5	p 153.9
3	R 15.3	D 15.8	M 360.9	R 145.7
4	K 18.6	R 17.7	K 283.0	
5	PEC (+)	K 15.6	C (−)	
6	P 11.6	PEC (+)	K 242.9	
7	K 13.2	P 11.8	L 221.1	
8	I 10.8	K 13.8	D 146.7	
9	L 11.3	I 12.0	T 114.8	
10	M 10.9	L 11.9	D 113.0	
11	R 11.6	M 12.0	C (−)	
12	PEC (+)	R 15.4	F 118.2	
13	K 9.8	PEC (+)	P 65.9	
14	R 15.2	K 11.9	T 48.0	
15	D 9.4	R 16.1	C (−)	
16	S 7.1	D 11.1	T 48.7	
17	D 12.8	S 8.1	C (−)	
18	PEC (+)	D 11.3	R 66.7	
19	L 6.0	PEC (+)	P 24.1	
20	A 4.7	L 6.9	S 16.4	
21	K 5.2	A 5.6	G 15.9	
22	PEC (+)	K 6.3	F 17.4	
23	T 3.8	PEC (+)	C (−)	
24	PEC (+)	T 3.7	G 11.4	
25	Q 3.9	PEC (+)		
26	E 1.7	Q 4.6		
27	S 4.3	E 1.9		
28	G 3.2	S 4.0		
29	Y 2.1	G 4.0		
30	PEC (+)	Y 2.4		
31	G 2.4	PEC (+)		
32		G 2.4		

Amino acids are reported as recovered pmoles (without correction), (+) indicates that the phenylthiohydantoin amino acid was detected as PTH-pyridylethyl cysteine (PTH-PEC) but not quantified; (−) indicates that the PTH-Cys was not detected, but it was deduced from the data for PE-SETI–V.

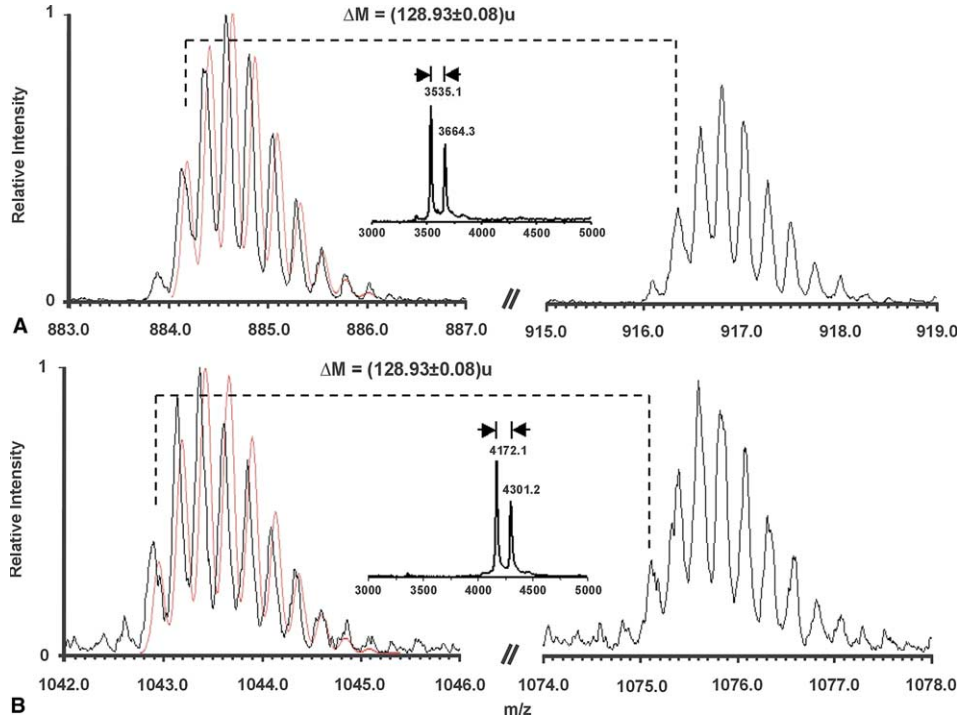

Fig. 2. Mass spectrometric analysis of SETI–II forms. ESI spectra of the quadruply charged state of components present in the unmodified SETI–II (panel A) and pyridylethylated SETI–II (PE-SETI–II) (panel B). Solid black lines correspond to experimental spectral traces; grey lines depict theoretical isotopic distributions expected for SETI–IIa forms. A0 peaks between the isotopic envelopes for SETI–IIa (left) and SETI–IIb (right) from the same preparation are linked by a dashed line and the mass difference calculated from the quadruply charged state is shown atop this line. MALDI-TOF spectra of the unmodified SETI–II (panel A, inset) and PE-SETI–II (panel B, inset) show two species, in each case separated by 129.1 ± 0.1 u. The expected average masses from the SETI–IIa sequence are 3535.1 u for the unmodified form and 4172.0 u for the pyridylethylated form. [See text for details.]

and 0 for SETI–IIb (Fig. 3). The quantitative sequencing data and the mass spectrometric measurements provides evidence for the structures proposed for SETI–IIa and SETI–IIb in Fig. 4.

2.3. SETI–V

The amino acid composition of SETI–V (Table 1) corresponded to 27 residues/mole before reduction and pyridylethylation and 24 residues/mole afterwards. One residue each of Cys, Pro and Arg was lost upon reduction and alkylation. When unmodified SETI-V was sequenced by Edman degradation (Table 2), one residue, PTH-Ile, was obtained in the first cycle, PTH-Leu and PTH-Pro were recovered in the second, and PTH-Met and PTH-Arg were released in the third. Thereafter, a single sequence corresponding to KCKLDTDCFPTCTCRPSGFCG was obtained (see Table 2). When PE-SETI-V was sequenced (data not shown), it did not have a double signal and started with Ile, Leu and Met, followed by KCKL. We hypothesized that the reactive site of the inhibitor had been cleaved during isolation with the trypsin affinity column. Thus, cleavage at the peptidyl bond Arg^3-Ile^4 would generate two polypeptides with sequences C^1PR^3 and I^4LM FCG^{27}, linked by a disulfide bond in the native preparation (arrow in Fig. 4). This peptide would account for the double signal observed in the sequencing experiments up to the third cycle. For PE-SETI-V, the absence of a double signal until the third cycle may be explained by the loss of the tripeptide C^1PR^3 after the reduction of the disulfide bond between the cysteines and the purification by HPLC.

Mass spectrometric results were consistent with the sequence proposed in Fig. 4 with measured average masses of 2990.5 u ($M_{Calc} = 2990.6$ u, $\Delta M = -0.1$ u) for the unmodified form, 2640.3 u ($M_{Calc} = 2640.2$ u, $\Delta M = +0.1$ u) for the chemically reduced form and 3166.0 u ($M_{Calc} = 3165.9$ u, $\Delta M = +0.1$ u) for the PE-SETI-V (spectra not shown). It should be pointed out that the mass decrease observed in the last two forms was mainly due to loss of the tripeptide C^1PR^3 after reduction, as previously explained. These results confirmed that a cleavage between residues Arg^3 and Ile^4 at the reactive site of the inhibitor had occurred probably during trypsin affinity chromatography.

2.4. Titration of inhibitors

The mixture of SETI–II a and b was assigned a molecular weight of 3535 Da which is the average of their molecular weights. An inhibitor-to-trypsin molar ratio of 1.07 was obtained from activity titration data (Fig. 5). The inhibitor-to-trypsin molar ratio was 1.12 when the same study was carried out with SETI-V using a molecular weight of 2997 Da. The linear titration curve that extrapolates to 1:1 molar stoichiometry and >95% inhibition is typical of strong trypsin inhibitors.

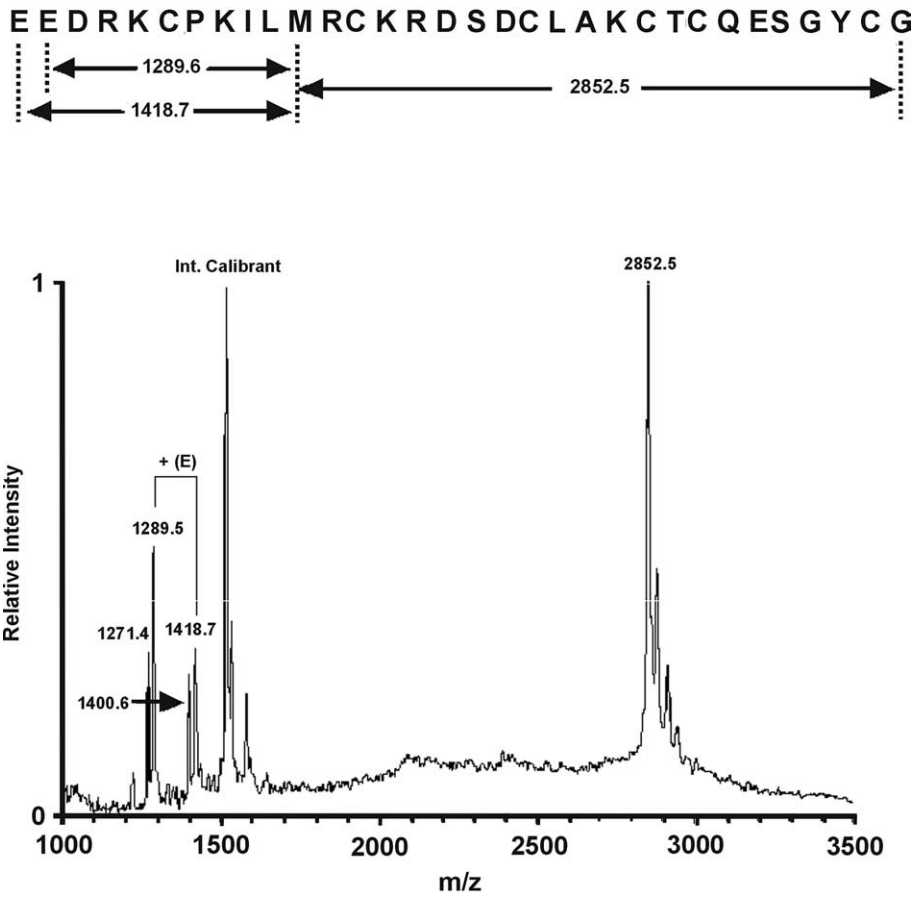

Fig. 3. MALDI-TOF/MS analysis of the cyanogen bromide cleavage products from reduced and pyridylethylated SETI–II (PE-SETI–II). Two N-terminal peptides are present at average masses of 1289.5 u (DM = −0.1 u) and 1418.7 u (DM = 0 u), and only a single carboxyl terminal peptide at mass 2852.5 u (DM = +0.1 u). The mass pairs (1418.7,1400.6) and (1289.5, 1271.4) correspond respectively to the open and closed forms of the β-lactam ring formed by modification of the side chain of the methionyl residue upon cleavage by CNBr (homoserine or homoserine lactone). Three internal calibrant peaks were used to calibrate the spectrum (only one is shown here). Peaks not identified in the spectrum correspond to sodium and copper adduct ions. [See text for experimental details and discussion.]

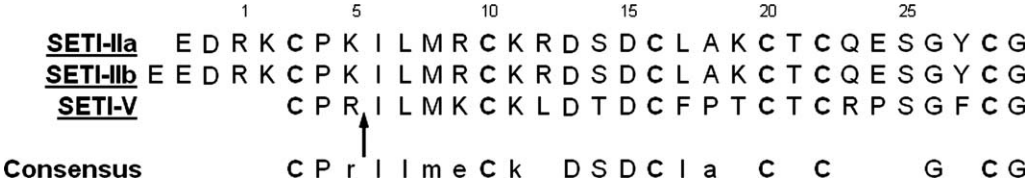

Fig. 4. Homology of the trypsin inhibitors from *Sechium edule*. The sequences for the three inhibitors were determined by automated Edman sequencing and confirmed by mass spectrometry. Cysteine was not identified in the native protein, but was identified positively when the reduced and pyridylethylated form of the inhibitor were sequenced. Data for the consensus sequence were obtained by aligning 37 squash-type trypsin inhibitors using the alignment tool Multalin (URL: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_server.html) from the Pôle Informatique Lyonnais, France. Invariant residues are printed in capital letters and the most preserved residues in many inhibitors in lower case letters in the consensus sequence.

2.5. Dissociation constants of the inhibitors

The dissociation constants were calculated from the same data (Fig. 5) by Morrison's procedure using the Gra Fit program (Knight, 1986) with K_i of 5.4×10^{-11} M and 1.12×10^{-9} M for SETI–II and SETI-V, respectively (see Section 3). The K_i values were within the range described for inhibitors of the "squash" family (10^{-8}–10^{-11} M) (Otlewski, 1990; Hamato et al., 1992).

2.6. Homology

Fig. 4 shows the primary structures of the inhibitors documented here and the consensus sequence obtained from 37 squash inhibitors. There is 29% identity of the amino acid sequence (not amino acid conserved) among the 37 inhibitors (algorithm Multalin from the Pôle Informatique Lyonnais, France, Network Protein Sequence Analysis, URL. http://npsa-pbil.ibcp.fr/cgi-bin/

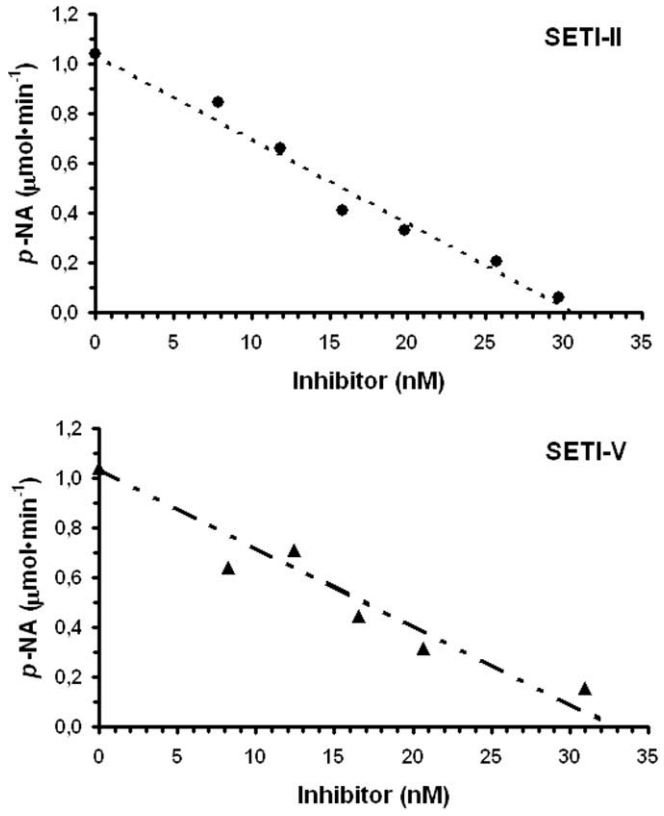

Fig. 5. Titration of bovine trypsin with SETI–II and SETI–V. The assay solution contained 0.28 mM DL-BAPNA and 28.6 nM trypsin in 50 mM Tris–HCl buffer, pH 8.2 with 20 mM calcium chloride. The inhibitor-to-trypsin molar ratios determined for SETI–II and SETI–V were 1.07 (top) and 1.12 (bottom), respectively.

npsa_automat.pl?page=/NPSA/npsa_server.html). The most striking feature of the sequence is the conservation of the number and position of the six cysteinyl residues. One would expect that the disulfide bridges should be the same in all inhibitors of this family. Indeed they are: 1–4, 2–5 and 3–6 as determined by NMR for EETI-II (Heitz et al., 1989), by crystallography for CMTI-I (Bode et al., 1989) and by wet chemistry for MCTI-A (Hara et al., 1989) and SETI–IIa (Faça et al., 2004).

The arrow in Fig. 4 indicates the position of the reactive center consisting of Lys′ or Arg′ and Ile″ using the notation of Schechter and Berger (1967). Variations in the amino acid sequence are concentrated in the N-terminus region, i.e., before Cysl. Four other residues are invariant in the sequence: Pro (P2), Ile (P1′), Gly (P22′) and Gly (P24′). All of these residues in CMTI-I are in contact with the trypsin protease (Bode et al., 1989). In general there is only one aromatic amino acid in each molecule, located at position seven (P2′) or 27 (P22′), except for SETI-V which contains an extra aromatic residue (Phe) at P12′.

Before the first invariant cysteine in SETI–II and SETI–V, other amino acids are present that extend the amino-terminus. If the inhibitor loses these N-terminal amino acids, for example a tripeptide, the association constant increases around 10-fold (Wieczorek et al., 1985). However, when an extra residue such as pyroglutamic acid is added to the amino terminus the association constant decreases 2.5-fold. A Lys residue at the P1 position inhibits trypsin 2.7-fold stronger than an Arg at the same position (Otlewski, 1993), and we observe a decrease in the association constant when SETI–II is compared to SETI–V (see Fig. 4).

The present study documents the isolation and the primary structure of a squash-type trypsin inhibitor. Its structure is identical to others described in the literature, especially the relative positions of the cysteines residues, reactive site and disulfide bridges (Faça et al., 2004). The loop which presents the reactive site (Arg or Lys (P1) and lie (P1′)) is conserved and provides the limited number of contacts with the reaction center of the trypsin, which leads to binding constants of 10^{-9}–10^{-10}.

3. Experimental

3.1. Plant material

S. edule (central city food market) was obtained from a commercial source. The seeds were collected and triturated using a domestic blender.

3.2. Inhibitor preparation

3.2.1. Extraction of inhibitors from S. edule and acetone precipitation

Fresh seeds (3.2 kg) obtained from *S. edule* were triturated in 100 mM ammonium acetate buffer (10 L, pH 7.0) and mixed-for 64 h at 4 °C. After centrifugation (5000*g* for 10 min at 4 °C) and lyophilization, the residue was suspended in distilled H_2O (800 ml), with the latter adjusted to aqueous Me_2CO (6:4, v/v) and then centrifuged (5000*g* for 10 min). The supernatant was next adjusted to aqueous Me_2CO (9:1, v/v) and centrifuged under the same conditions. A yellow liquid phase formed at the bottom of the tube containing 73% of total trypsin inhibitor activity was collected.

3.2.2. Sephadex G-25 size exclusion chromatography

A 20-ml aliquot of the yellow liquid obtained by acetone fractionation (353 mg protein and 194 IU) was diluted to 50 ml with 5% aqueous acetic acid (HOAc, v/v) and applied to a Sephadex G-25 column (2.5 × 250 cm). The column, previously standardized as described by Freitas et al. (1993), was eluted with 5% aqueous HOAc (v/v) at 84 ml/h at room temperature with 14-ml fractions were collected. The effluent was monitored at 280 nm and trypsin inhibitor activity was measured with DL-BAPNA as described below.

3.2.3. Trypsin-Sepharose affinity chromatography

Trypsin-Sepharose-4B resin (Pharmacia) was prepared by the method of Cuatrecasas (1970) using bovine trypsin. An aliquot of the eluate from the G-25 column (55.3 IU

and 14.1 mg protein) in 5% aqueous HOAc (750 μl) was applied to 8 ml of a trypsin-Sepharose resin bed equilibrated with 100 mM Tris–HCl, pH 8.0. Protein bound to the resin was eluted with 56 ml 500 mM KCl, pH 2.0. Fractions (4 ml) were collected into tubes containing 200 μl 1 M Tris–HCl, pH 8.0.

3.2.4. Purification by RP-HPLC

Protein eluted from the trypsin-Sepharose affinity column (0.23 mg protein containing 5.7 IU) was diluted to 470 μl with 0.1% TFA and loaded on a Vydac C_{18} column (4.6 mm i.d. × 250 cm), equilibrated for 20 min with 90% A and 10% B (solution A: 0.1% aqueous TFA; solution B: 0.09% TFA in MeCN2–H_2O (4:1)). After five minutes of isocratic elution with 90% A, a linear gradient of 1% B per minute was applied to the column for 33 min at 1 ml/min at room temperature. The eluate was monitored spectrophotometrically at 220 and 280 nm simultaneously and peaks were collected manually.

3.2.5. Reduction and alkylation

The inhibitor (containing 6 moles cysteine/mole inhibitor by amino acid analysis) was reduced and alkylated with 4-vinylpyridine (Friedman et al., 1970; modified by Meneghatti et al., 1992). Briefly 1 mole cysteine was reacted with 29 moles DTT, in 250 mM Tris–HCl buffer, pH 8.2, under nitrogen for 4 h at 50 °C in the dark. Then, 52 moles 4-vinylpyridine were added and the reaction was left to stand under nitrogen for an additional 4 h in the dark at room temperature. The reaction was stopped by freezing in dry ice. PE-inhibitor was purified by RP-HPLC under the same conditions as used for the unmodified inhibitor.

3.2.6. Trypsin activity

The trypsin used in these studies contained 55.4% active trypsin (a stock solution was titrated with *p*-nitrophenyl guanidine benzoate chloride (NPGB) by the method of Chase and Shaw (1967)). All concentrations of trypsin reported here are for active trypsin. Trypsin activity was determined by the method of Erlanger et al. (1961) using benzoyl DL-arginyl-*p*-nitroanilide (DL-BAPNA) as substrate. A solution of active trypsin containing 13.60 μg/ml in 50 mM Tris–HCl buffer, pH 8.2, with 20 mM $CaCl_2$ (referred to as Tris calcium buffer = TCB) and a solution of 0.3 mM DL-BAPNA in TCB were prepared daily.

The inhibitor in 50 μl TCB was preincubated with 50 μl active trypsin (13.60 μg/ml) in a 1 ml cuvette for 5 min at 37 °C. The reaction was started by the addition of 900 μl 0.3 mM DL-BAPNA in TCB and the product, *p*-nitroanilide (*p*-NA, with 8800 M^{-1} cm^{-1} molar absorptivity Erlanger et al., 1961), was monitored at 410 nm, and recorded at 0.1 AUFS for 4–6 min. Each measurement was made in triplicate and differences were ⩽5%.

Trypsin activity is reported as μmoles *p*-NA hydrolyzed per min (μmol/min), with inhibitor activity (IU) reported in the same units. Specific activity is reported as activity or inhibition per mg (or μg) protein. The trypsin used in this study had a specific activity of 1.54 μmoles *p*-NA/μg active trypsin.

3.2.7. Determination of the dissociation constant and stoichiometry

The titration curve of trypsin with varying amounts of inhibitor was used to determine the dissociation constants and the stoichiometry for each inhibitor. Active trypsin (50 μl, 0.571 μM) and inhibitor (50 μl, 0–0.650 μM) were preincubated in TCB for 5 min at 37 °C. The trypsin assay was started by adding 900 μl 0.3 mM DL-BAPNA in TCB. The final concentrations were 28.6 nM active trypsin, 0–32 nM inhibitor and 0.270 mM DL-BAPNA. Residual trypsin activity was measured by release of *p*-NA that was monitored at 410 nm for 4–6 min at 0.1 AUFS. Each measurement was made in triplicate and differences were <5%.

The K_i was calculated using Morrison's equation with the Gra Fit program (Knight, 1986).

Morrison's equation:

$$\frac{v_i}{v_0} = \frac{E_0 - I_0 - K_i + \sqrt{(E_0 + I_0 - K_i)^2 - 4E_0I_0}}{2E_0}$$

where v_0 = velocity without inhibitor; v_i = velocity with the inhibitor; E_0 = total concentration of the enzyme; I_0 = total concentration of the inhibitor; K_i = inhibition constant.

The stoichiometry of the reaction was obtained graphically by extrapolation of the linear regression line plot of v_i/v_0 versus inhibitor concentration.

3.2.8. Amino acid composition and protein determination

Amino acid compositions and total amount of protein were determined by amino acid analysis using the phenylthiocarbamyl (PTC) derivatization method (Bidlingmeyer et al., 1984). Briefly, samples were hydrolyzed with 6 N HCl in the vapor phase (110 °C) for 22 h (Atherton, 1989), then derivatized with phenylisothiocyanate (PITC, Pierce), and applied to an RP-HPLC Picotag column. An external standard (Amino Acid Standard H, Pierce Chemical Co., Rockford, IL), 100 pmoles, was used to calculate the amino acid content of the sample.

3.2.9. Edman degradation

The inhibitors were sequenced with a Procise 491 (Applied Biosystem, Foster City, CA) sequencer using a fiberglass membrane treated with polybrene and a pulse liquid program. The HPLC system was calibrated with 10 pmoles of the standard mixture of phenylthiohydantoin (PTH) amino acids (Applied Biosystem).

3.2.10. CNBr cleavage

Inhibitor, 50 pmoles in H_2O–MeOH–HCO_2H (5:3:2, v/v/v) containing a 500-fold molar excess of CNBr (Pierce Chemical Co.) to methionyl residues, was allowed to react for 30 min in the dark at room temperature. The reaction mixture was analyzed by MALDI-TOF/MS immediately after CNBr treatment.

3.2.11. MALDI-TOF/MS

Samples were dissolved in a saturated solution of α-cyano-4-hydroxy-cinnamic acid prepared in H_2O–MeCN (2:1, v/v) and 0.1% TFA to a final concentration of 400 fmol/μl. Aliquots of 0.5 μl (200 fmol) were deposited on gold-coated sample plates and allowed to air-dry. The dry spots were washed twice with 0.5 μl of cold 0.1% aqueous TFA solution, excess liquid being removed by vacuum suction, and air-dried before the sample plate was introduced into the ionization source of the mass spectrometer. Two MALDI-TOF mass spectrometers were used in this study. The principal device was a commercial instrument, STR Voyager (model STR, PE Biosystems, Foster City, CA) with delayed extraction and a reflector. This instrument is equipped with a nitrogen laser that delivers pulses of ultraviolet light of 337 nm at 2 Hz to the matrix spots, with each pulse yielding a full mass spectrum. One-hundred individual shots were averaged with software provided by the manufacturer using 0.5 ns data channel widths to produce a single mass spectrum.

The other instrument (Beavis and Chait, 1989, 1990) was constructed in the Laboratory for Mass Spectrometry and Gaseous Ion Chemistry at the Rockefeller University and details of its operation can be found elsewhere (Cohen and Chait, 1996). This device is equipped with a Neodymium-YAG laser which pulses UV light at 355 nm. Data collection and analysis were performed with in-house software. One-hundred individual shots were averaged to produce a single spectrum.

3.2.12. ESI/MS

ESI/MS analysis was performed on two different instruments, each fitted with an electrospray ionization source built at the Rockefeller University: an ion trap mass spectrometer, model LCQ (ThermoQuest Finnigan, San Jose, CA) and a triple quadrupole mass spectrometer, model TSQ-700 (ThermoQuest Finnigan, San Jose, CA). The instruments were previously calibrated and tuned according to manufacturer's instructions. Samples for ESI/MS analysis were prepared by dilution to a final concentration range of 20–200 fmol/μl in a mixture of H_2O–MeOH–HOAc at 49:50:1 (v/v/v), and infused at constant flow rates of 0.5 or 1.0 μl/min with an infusion pump, model 22 (Harvard, South Natick, MA), through a 50-μm i.d. fused silica capillary directly into the ionization source of the mass spectrometer. Desolvation conditions for peptide ions included maintaining the heated capillary at 150 °C and declustering potentials of at least +38 V. Data were collected in the profile mode and 100 spectra were averaged during acquisition time to produce a single final spectrum.

Acknowledgements

HJL was the recipient of a graduate student fellowship from CNPq. JCP received a postdoctoral fellowship from FAPESP. Research was supported by FAPESP, CNPq and FINEP (PADCT). We thank Dr. Maria Luiza Vilela Oliva (Universidade Federal de São Paulo) for permitting us to prepare the trypsin affinity column in her laboratory and for many useful discussions, and Mrs. Elettra Greene for correcting the manuscript.

References

Atherton, D., 1989. Successful PTC amino acid analysis at the picomole level. In: Hugli, T.E. (Ed.), Techniques in Protein Chemistry. Academic Press Inc., San Diego, pp. 273–283.

Beavis, R.C., Chait, B.T., 1989. Factors affecting the ultraviolet laser desorption of proteins. Rapid Commun. Mass Spectrom. 3, 233–237.

Beavis, R.C., Chait, B.T., 1990. High-accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry. Anal. Chem. 62, 1836–1840.

Bidlingmeyer, B.A., Cohen, S.A., Tarvin, T.L., 1984. Rapid analysis of amino acids using pre-column derivatization. J. Chromatogr. (Biomed. Appln.) 336, 93–104.

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

Bode, W., Huber, R., 1993. Structural basis of the proteinase–protein inhibitor interaction. In: Avilés, F.X. (Ed.), Innovations in Proteases and their Inhibitors. Walter de Gruyter, Berlin, pp. 81–122.

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

Chase Jr., T., Shaw, E., 1967. *p*-Nitrophenyl-*p*′-guanidinobenzoate HCl: A new active site titrant for trypsin. Biochem. Biophys. Res. Commun. 29, 508–514.

Cohen, S.L., Chait, B.T., 1996. Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. Anal. Chem. 68, 31–37.

Cuatrecasas, P., 1970. Protein purification by affinity chromatography. Derivatizations of agarose and polyacrylamide beads. J. Biol. Chem. 245, 3059–3065.

Erlanger, B.F., Kokowsky, N., Cohen, W., 1961. The preparation and properties of two new chromogenic substrates of trypsin. Arch. Biochem. Biophys. 95, 271–278.

Faça, V.M., Pereira, S.R., Laure, H.J., Greene, L.J., 2004. Determination and reoxidation of the disulfide bridges of a squash-type trypsin inhibitor from *Sechium edule* seeds. Protein J. 23, 309–315.

Freitas, O., Padovan, G.J., Vilela, L., dos Santos, J.E., de Oliveira, J.E.D., Greene, L.J., 1993. Characterization of protein hydrolysates prepared for enteral nutrition. J. Agric. Food Chem. 41, 1432–1438.

Friedman, M., Krull, L.H., Cavins, J.F., 1970. The chromatographic determination of cystine and cysteine residues in proteins as *S*-β-(4-pyridylethyl)cysteine. J. Biol. Chem. 245, 3868–3871.

Hamato, N., Takano, R., Kamei-Hayashi, K., Hara, S., 1992. Purification and characterization of serine proteinase inhibitors from gourd (*Lagenaria leucantha* Rusby var. *Gourda* Makino) seeds. Biosci., Biotechnol., Biochem. 56, 275–279.

Hara, S., Makino, J., Ikenaka, T., 1989. Amino acid sequences and disulfide bridges of serine proteinase inhibitors from bitter gourd (*Momordica charantia* LINN.) seeds. J. Biochem. 105, 88–92.

Heitz, A., Chiche, L., Le-Nguyen, D., Castro, B., 1989. ^{1}H 2D NMR and distance geometry study of the folding of *Ecballium elaterium* trypsin inhibitor, a member of the squash inhibitors family. Biochemistry 28, 2392–2398.

Hider, R.C., Drake, A.F., Morrison, I.E.G., Kupryszewski, G., Wilusz, T., 1987. Structure analysis of inhibitors isolated from Cucurbitacae seeds – (Circular dichroism studies). Int. J. Peptide Protein Res. 30, 397–403.

Knight, C.G., 1986. The characterization of enzyme inhibition. In: Barrett, A.J., Salvesen, G. (Eds.), Proteinase Inhibitors. Elsevier Science, Amsterdam, The Netherlands, pp. 23–51.

Krishnamoorthi, R., Gong, Y., Richardson, M., 1990. A new protein inhibitor of trypsin and activated Hageman factor from. FEBS Lett. 273, 163–167.

Laskowski Jr., M., Kato, I., 1980. Protein inhibitors of proteinases. Ann. Rev. Biochem. 49, 593–626.

Le-Nguyen, D., Mattras, H., Coletti-Previero, M-A., Castro, B., 1989. Design and chemical synthesis of a 32 residue chimeric microprotein inhibiting both trypsin and carboxypeptidase A. Biochem. Biophys. Res. Commun. 162, 1425–1430.

McQuade, T.J., Tomasselli, A.G., Lui, L., Karacostas, V., Moss, B., Sawer, T.K., Heinrikson, R.L., Tarpley, W.G., 1990. A synthetic HIV-1 protease inhibitor with antiviral activity arrests HIV-like particle maturation. Science 247, 454–456.

Menegatti, E., Tedeschi, G., Ronchi, S., Bortolotti, F., Ascenzi, P., Thomas, R.M., Bolognesi, M., Palmieri, S., 1992. Purification, inhibitory properties and amino acid sequence of a new serine proteinase inhibitor from white mustard (*Sinapis alba* L.) seed. FEBS Lett. 301, 10–14.

Otlewski, J., 1990. The squash inhibitor family of serine proteinases. Biol. Chem. Hoppe-Seyler 371 (Suppl.), 23–28.

Otlewski, J., 1993. The squash inhibitors of serine proteinases. In: Avilés, F.X. (Ed.), Innovations in Proteases and their Inhibitors. Walter de Gruyter, Berlin, pp. 369–388.

Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Wilusz, T., Polanowski, A., 1989. Synthesis of an elastase inhibitor by monosubstitution of arginine-5 with valine at the reactive site in a trypsin inhibitor from squash seeds (CMTI III). Biol. Chem. Hoppe-Seyler 370, 499–502.

Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., 1991. Chemical synthesis of new trypsin, chymotrypsin and elastase inhibitors by amino-acid substitutions in a trypsin inhibitor from squash seeds (CMTI III). Biol. Chem. Hoppe-Seyler 372, 63–68.

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

Wieczorek, M., Otlewski, J., Cook, J., Parks, K., Lulek, J., Wilimowska-Pelc, A., Polanowski, A., Wilusz, T., Laskowski Jr., M., 1985. The squash family of serine proteinase inhibitors. Amino acid sequences and association equilibrium constants of inhibitors from squash, summer squash, zucchini, and cucumber seeds. Biochem. Biophys. Res. Commun. 126, 646–652.