J. Biochem. 119, 1106-1113 (1996)

Purification, Characterization, and Sequencing of Two Cysteine Proteinase Inhibitors, Sca and Scb, from Sunflower (*Helianthus annuus*) Seeds

Yoshiaki Kouzuma,* Keiko Kawano,* Makoto Kimura,*,[1] Nobuyuki Yamasaki,* Tomoko Kadowaki,† and Kenji Yamamoto†

*Laboratory of Biochemistry, Faculty of Agriculture, and †Department of Pharmacology, Faculty of Dentistry, Kyushu University, Higashi-ku, Fukuoka 812

Received for publication, January 12, 1996

Two proteinaceous cysteine proteinase inhibitors (cystatins) referred to as Sca and Scb were purified to homogeneity from the seeds of sunflower (*Heliantas annuus*) by gel filtration on Sephadex G-75 followed by a series of ion-exchange column chromatographies and reverse-phase HPLC (RP-HPLC). The isoelectric points (pI) of Sca and Scb were estimated to be 5.6 and 9.6, respectively. The inhibitory potencies of these two cystatins were examined with cysteine proteinases from various sources, such as plants, mammals, and bacteria. Papain was strongly inhibited by both Sca and Scb with K_i values of 5.6×10^{-9} and 1.7×10^{-10} M, respectively. Sca and Scb were also found to be potent inhibitors of ficin (K_i values of 1.9×10^{-6} and 2.8×10^{-9} M, respectively). Rat cathepsin H was inhibited strongly by Scb and slightly by Sca. Although rat cathepsins B and L were significantly inhibited by Scb, they were scarcely affected by Sca. Neither Sca nor Scb inhibited Arg-gingipain, an arginine-specific cysteine proteinase of *Porphyromonas gingivalis*. The complete amino acid sequences of the two inhibitors were determined by protein chemical methods. The proteins Sca and Scb consist of 83 and 101 amino acid residues with M_r of 9,330 and 11,187, respectively, and there are identical residues at 34 positions in the two sequences, that is at 42% of the residues compared. Comparison of their sequences with those of other cystatins revealed that Sca shares 59-73% identical residues with other phytocystatins, while Scb shows less identity to other phytocystatins, sharing only 28-38% identical residues. Furthermore, only 20-27% of the residues of both cystatins, Sca and Scb, are identical to those of the animal cystatins.

Key words: amino acid sequence, cysteine proteinase inhibitor, *Helianthus annuus*, phytocystatin, sunflower.

Proteinaceous cysteine proteinase inhibitors have been found in various animal tissues and fluids, and many of them have been isolated and characterized in terms of protein structure and inhibitory activity. These studies revealed that, despite the great diversity of protein structures, cysteine proteinase inhibitors have many properties in common and they comprise the cystatin superfamily, which can be derived into three families (family 1-family 3) (for a review see Ref. *1*). In addition to animal cystatins, oryzacystatin I, which is the best characterized cysteine proteinase inhibitor from a plant, was first isolated from rice seeds, and its inhibitory activity and cDNA structure were investigated (*2-4*). Subsequently, information on the amino acid sequences of some other cysteine proteinase inhibitors from plants, including oryzacystatin II (*5*), cowpea cystatin (*6*), corn cystatin (*7*), and avocado cystatin (*8*) has become available either by deduction from cDNA sequences or by direct sequencing of the protein. Comparison of the amino acid sequences of plant cystatins with those of animal cystatins showed that although plant proteins are homologous to family 1 cystatins with respect to lack of disulfide bonds, their sequences are more homologous to those of family 2 cystatins than those of family 1 cystatins. Thus, it has been generally accepted that plant cystatins should be classified into a fourth cystatin family, "phytocystatin" (*5*).

In addition to phytocystatins, three cysteine proteinase inhibitors, whose amino acid sequences are similar to those of the Kunitz family proteinase inhibitors, were isolated from potato tubers (*9, 10*), and it was reported that they strongly inhibited cathepsin L, but exhibited only weak inhibitory activity toward papain (*9*).

Previously, we found inhibitory activity toward papain in the protein extract from sunflower seeds. To get more structural and functional information about cysteine proteinase inhibitors of plant origin, we attempted to purify cystatin from sunflower seeds. In this paper, we report on the purification of two cysteine proteinase inhibitors (des-

[1] To whom correspondence should be addressed. E-mail: mkimura@agr.kyushu-u.ac.jp

Abbreviations: CNBr, cyanogen bromide; cystatin, cysteine proteinase inhibitor; DMSO, dimethyl sulfoxide; IBA; *o*-iodosobenzoic acid; pI, isoelectric point; PITC, phenyl isothiocyanate; RP-HPLC, reverse-phase high-performance liquid chromatography; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid.

ignated as Sca and Scb), and characterization of their inhibitory activities toward cysteine proteinases. Furthermore, we describe the complete amino acid sequences of Sca and Scb, and discuss their structural relationship to other phytocystatins as well as to animal cystatins.

MATERIALS AND METHODS

Materials—Sunflower (*Helianthus annuus*) seeds were purchased from Takii Seed (Kyoto). Trypsin, chymotrypsin, papain, ficin, and bromelain were obtained from Sigma Chemical (St. Louis). Lysylendopeptidase, *Staphylococcus aureus* V8 protease, and reagents for protein sequencing were purchased from Wako Pure Chemicals (Osaka). L-Pyroglutamyl-L-phenylalanyl-L-leucine *p*-nitroanilide (Pyr-Leu-Phe-*p*NA), carbobenzoxy-L-phenylalanyl-L-arginine 4-methyl-7-coumarylamide (Z-Phe-Arg-MCA), *t*-butyloxycarbonyl-L-phenylalanyl-L-seryl-L-arginine 4-methyl-7-coumarylamide (Boc-Phe-Ser-Arg-MCA), and L-arginine 4-methyl-7-coumarylamide (Arg-MCA) were obtained from Peptide Institute (Minoh). Cathepsins B and H were purified from rat spleen as described previously (*11, 12*), with a slight modification (*13*). Rat kidney cathepsin L was purified to homogeneity essentially as described by Bando *et al.* (*14*). The purification of Arg-gingipain from *Porphyromonas gingivalis* was performed as described previously (*15*). Sephadex G-75, S-Sepharose, and Q-Sepharose were from Pharmacia Biotech (Uppsala, Sweden), and ampholine (pH 3.5–10.0) was from LKB (Bromma, Sweden). DEAE-cellulose (DE52) and CM-cellulose (CM52) were from Whatman BioSystems (Maidstone, Kent, England). All other chemicals were of analytical grade for biochemical use.

Purification of Sunflower Cystatins—Sunflower cystatins Sca and Scb were purified as follows. Husked sunflower seeds (400 g) were ground and defatted with petroleum ether. The defatted meal (200 g) was suspended in 500 ml of 10 mM phosphate buffer saline, pH 7.2, and stirred at 4°C for 2 h. After centrifugation, the extract was salted out by saturation with ammonium sulfate. The precipitate was recovered, dissolved, and dialyzed against 10 mM phosphate buffer, pH 7.2. The crude inhibitor thus obtained was chromatographed on Sephadex G-75 (2.8×150 cm) with the same buffer. The fraction exhibiting the inhibitory activity toward papain was then applied to a DEAE cellulose column (2.8×35 cm) equilibrated with 50 mM Tris-HCl buffer, pH 8.5, and the protein was eluted with a linear gradient of increasing NaCl concentration from 0 to 0.3 M in the buffer.

The protein (designated as Sca) adsorbed on DEAE cellulose was applied to a Q-Sepharose column (1.5×15 cm) equilibrated with 50 mM Tris-HCl buffer, pH 8.7, and the protein was eluted with a linear gradient of increasing NaCl concentration from 0 to 0.3 M. The protein (designated as Scb) unadsorbed on DEAE cellulose was applied to a CM-cellulose column (2×35 cm) equilibrated with 10 mM phosphate buffer, pH 7.2 and eluted with a linear gradient of increasing NaCl concentration from 0 to 0.2 M. The protein Scb was further purified on S-Sepharose (1.5×15 cm) with the same buffer as that for CM-cellulose column chromatography.

Both cystatins were finally purified by RP-HPLC on a YMC GEL C_4 column (4.6×250 mm) equilibrated with 0.1% trifluoroacetic acid (TFA) and chromatographed using a linear gradient from 0 to 56% acetonitrile in 0.1% TFA over 30 min at flow rate of 1.0 ml/min. The elute was monitored by measuring the absorbance at 220 nm, and each peak was collected manually.

The purity of a protein was judged by tricine SDS-PAGE, which was carried out on a 15% polyacrylamide gel, as described by Schagger and Jagow (*16*). The isoelectric point of a protein was measured by electrofocusing on a 5% polyacrylamide gel (0.5×9 cm) containing 2% carrier ampholyte with a pH range from 3.5 to 10.0 by the method of Svensson (*17*).

Assay of Proteinase Inhibitory Activity—The inhibitory activity toward papain and ficin activity was measured by the method of Filippova *et al.* (*18*). Briefly, papain was dissolved at 0.1 mg/ml in 100 mM phosphate buffer, pH 6.5, containing 0.3 M KCl, 0.1 mM EDTA and 1 mM DTT. One hundred microliters of enzyme solution was preincubated at 37°C for 10 min with 100 μl of various concentrations of the inhibitor solution. One milliliter of substrate solution (Pyr-Phe-Leu-pNA dissolved in the same buffer in 10% DMSO) was added at a final concentration of 420 μM and incubated at 37°C for 10 min. After incubation, 150 μl of 30% acetic acid was added and the absorbance was measured at 420 nm. Inhibitory activity was defined as the percentage inhibition of Pyr-Phe-Leu-pNA hydrolyzing activity of papain in the assay mixture. Purified cathepsins B, H, and L samples in 20 mM sodium phosphate buffer, pH 6.0, containing 10 mM dithiothreitol and 1 mM EDTA (0.5 ml) were preincubated at 37°C for 10 min with or without the proteinase inhibitors, before addition of 0.5 ml of the respective substrate solutions (Z-Phe-Arg-MCA for cathepsins B and L; Arg-MCA for cathepsin H) to start the reaction. The reaction was carried out at 40°C for 10 min and stopped by addition of 1 ml of 100 mM sodium acetate buffer, pH 5.0, containing 10 mM iodoacetic acid. The liberated product, 7-amino-4-methylcoumarin, was estimated by measurement of the fluorescence intensity at 460 nm using a Hitachi fluorescence spectrophotometer, model F-3010. The excitation wavelength was 380 nm. The inhibitory effects of Sca and Scb on Arg-gingipain were tested by preincubation at 37°C for 5 min followed by assay under the standard conditions with Boc-Phe-Ser-Arg-MCA as a substrate (*15*).

Sequence Determinations of Sca and Scb—Tryptic and lysylendopeptidase digestions were performed in 0.2 M *N*-methylmorpholine acetate buffer, pH 8.1, at 37°C for 6 h using an enzyme-to-substrate ratio of 1 : 50. *S. aureus* V8 protease digestion was done in 0.1 M ammonium acetate buffer, pH 4.0, using an enzyme/substrate ratio of 1 : 50. Cyanogen bromide (CNBr) (*19*) and *o*-iodosobenzoic acid (IBA) (*20*) cleavages were carried out in 70% formic acid and 80% acetic acid containing 8 M guanidine-HCl, *p*-cresol, and 1 mM EDTA, respectively. Peptide mixture obtained by enzymatic and chemical cleavages were separated by RP-HPLC on a YMC-GEL C_4 or Wakosil C_{18} column equilibrated with 0.1% TFA and eluted with a linear gradient of 80% acetonitrile in 0.1% TFA (flow rate: 0.6 ml/min). Peptides were hydrolyzed in the vapor of 5.7 N HCl containing 0.02% 2-mercaptoethanol at 110°C for 24 h in evacuated sealed tubes. Amino acid composition of each peptide was analyzed on a Shimadzu LC-6A system amino acid analyzer after derivatization of amino acids with PITC,

as described by Heinrikson and Meredith (*21*). Cysteine residues were evaluated as a carboxymethylated cysteine (CM-cystein) after reduction and carboxymethylation of the protein (*22*). Amino acids were sequenced by a gas-phase sequencer PSQ-1 (Shimadzu).

RESULTS

Purification of Sca and Scb—The crude sunflower cystatin obtained by saturation with ammonium sulfate was first chromatographed on Sephadex G-75. The fractions containing the inhibitory activity toward papain were pooled and concentrated by salting out with ammonium sulfate, and dialyzed against 50 mM Tris-HCl, pH 8.5. The protein solution was then applied to a column of DEAE-cellulose equilibrated with 50 mM Tris-HCl, pH 8.5, and eluted with a linear gradient of NaCl. As shown in Fig. 1, there were two main peaks with inhibitory activity toward papain. The protein eluted in the unadsorbed fractions (designated as Scb) was further purified by ion-exchange column chromatography on CM-cellulose followed by S-Sepharose in 10 mM phosphate buffer, pH 7.2 (Fig. 2), while the protein in the adsorbed fraction (designated as Sca) was further purified by ion-exchange column chromatography on Q-Sepharose in 50 mM Tris-HCl, pH 8.5, (Fig. 3). The final step in the purifications of Sca and Scb was performed by RP-HPLC on a YMC-GEL C_4 column in 0.1% TFA, as described in "MATERIALS AND METHODS" (Fig. 4). The yields of Sca and Scb were calculated to be about 0.3 and 0.5 mg, respectively, from 100 g of the seeds.

The purities of the proteins Sca and Scb thus obtained were analyzed by tricine SDS-PAGE. The analysis showed that each peak contained a pure protein, and that the molecular weights of Sca and Scb were 8.6 and 10.7 kDa, respectively (Fig. 4C). Furthermore, the pI values of Sca and Scb were estimated to be 5.6 and 9.6 by means of electrofocusing on a 5% polyacrylamide gel column containing 2% carrier ampholyte of pH 3.0-10.5 (data not shown).

Table I shows the amino acid compositions of the two sunflower cystatins, calculated on the basis of their molecular weights. Sca and Scb were composed of about 80 and 100

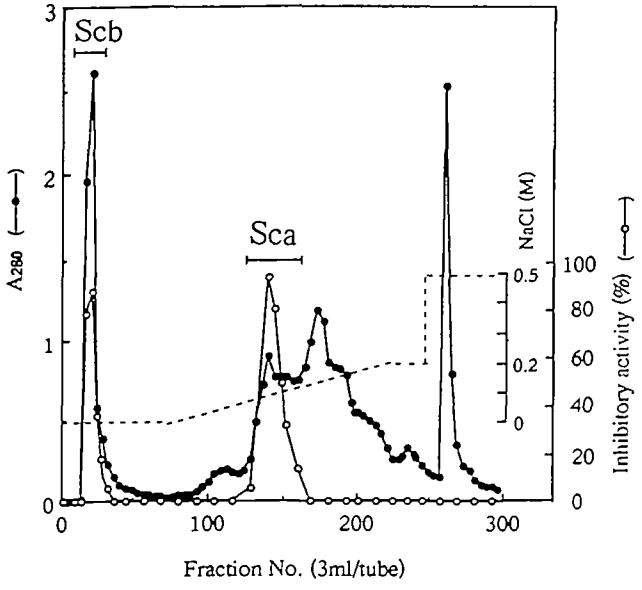

Fig. 1. **Ion-exchange column chromatography of the cysteine proteinase inhibitors on a DEAE cellulose column (2.8×35 cm).** Total proteins obtained from the sunflower seeds were first applied to a Sephadex G-75 column (2.0×150 cm) equilibrated with 10 mM Na-phosphate buffer, pH 7.2. Fractions containing the inhibitory activity against papain were concentrated by salting out and dialyzed against 50 mM Tris-HCl buffer, pH 8.5, and then applied to a DEAE cellulose column previously equilibrated with the same buffer. After the column was washed, the bound proteins were eluted with a linear gradient of NaCl from 0 to 0.3 M in the same buffer. Every second fraction was assayed for inhibitory activity against papain, as described in "MATERIALS AND METHODS." The active fractions (designated as Sca and Scb) were pooled as indicated by the horizontal bars.

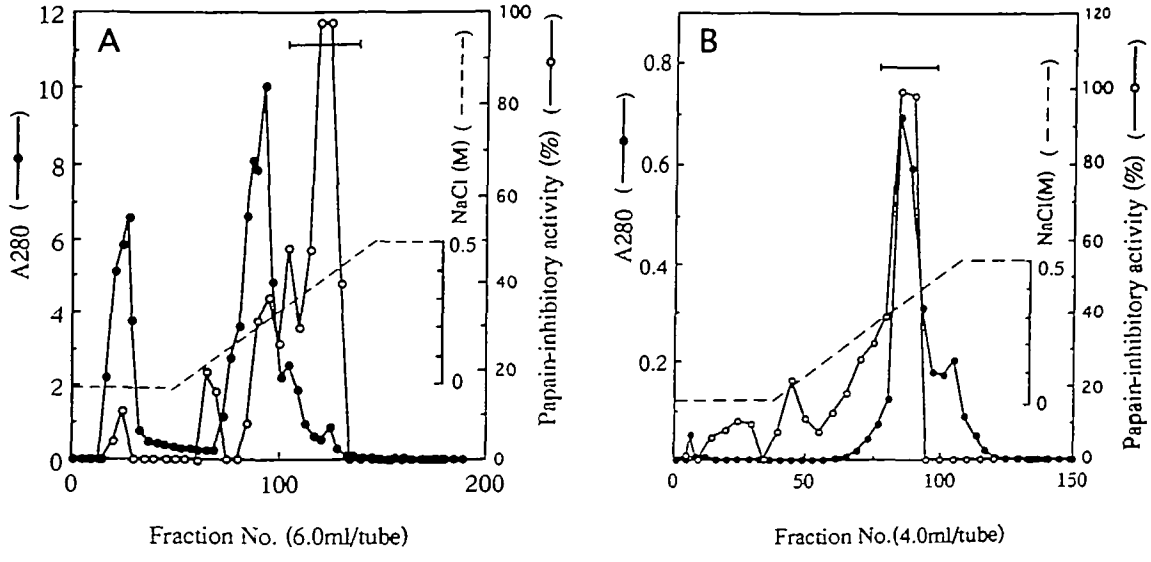

Fig. 2. **Purification of the sunflower cysteine proteinase inhibitor Scb by ion-exchange column chromatographies.** A: The unadsorbed fractions on the DEAE cellulose column were salted out, dialyzed against 10 mM Na-phosphate buffer, pH 7.2, and then purified on a CM-cellulose column (2×35 cm) as described in "MATERIALS AND METHODS." B: The fractions exhibiting the inhibitory activity were further purified on an S-Sepharose column (1.5×15 cm) in 10 mM phosphate buffer, pH 7.2.

amino acid residues, respectively. The analysis showed that cysteine residue was absent in both proteins. Sca and Scb include similar numbers of aspartic acid and glutamic acid residues, while Scb contains more arginine and lysine residues than Sca, in accordance with the estimated pI values of the two proteins.

Inhibitory Activities of Cystatins Sca and Scb—The inhibitory activities of two cystatins Sca and Scb were first investigated using papain and ficin: cysteine proteinases of plant origin. Figure 5 shows the titration patterns of papain with both sunflower cystatins at 10^{-7} M enzyme concentration. Although Sca and Scb could inhibit papain with 1 : 1 stoichiometry, their inhibitory activities were quite different: Scb inhibits papain more strongly than Sca in terms of the inhibitory constants. From this titration, the inhibitory constants (K_i values) for Sca and Scb for papain were calculated to be 5.6×10^{-9} and 1.7×10^{-10} M, respectively. Similarly, the inhibitory constants for Sca and Scb toward

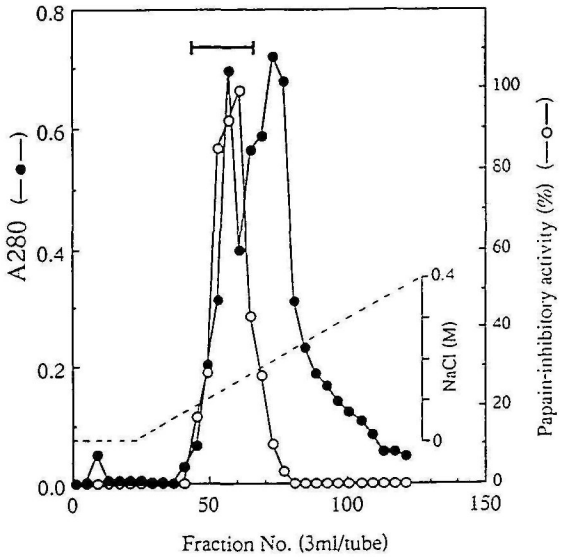

Fig. 3. **Purification of the sunflower cysteine proteinase inhibitor Sca on a Q-Sepharose column (1.5×15 cm).** The protein adsorbed on the DEAE-cellulose column was concentrated by salting out, dialyzed against 50 mM Tris-HCl buffer, pH 8.7, and then further purified on a Q-Sepharose column, as described in "MATERIALS AND METHODS."

TABLE I. **Amino acid compositions of Sca and Scb.**

Amino acid	Sca	Scb
Cys[a]	0.00 (0)	0.00 (0)
Asp	7.15 (7)	7.16 (7)
Thr	4.22 (4)	6.76 (7)
Ser	3.90 (4)	3.58 (4)
Glu	11.53 (12)	12.57 (13)
Pro	4.35 (4)	5.01 (5)
Gly	6.74 (7)	9.44 (9)
Ala	7.00 (7)	7.12 (7)
Val	5.27 (5)	10.03 (10)
Met	0.79 (1)	0.69 (1)
Ile	1.90 (2)	4.09 (4)
Leu	5.39 (5)	5.92 (6)
Tyr	2.72 (3)	3.67 (4)
Phe	3.41 (3)	2.90 (3)
Lys	7.54 (8)	14.25 (14)
His	1.00 (1)	1.27 (1)
Arg	2.10 (2)	3.05 (3)
Trp[b]	2.03 (2)	1.00 (1)
Total	77	98
MW	8,610	10,822

[a]Determined as *S*-carboxymethylated cysteine. [b]Estimated spectrophotometrically.

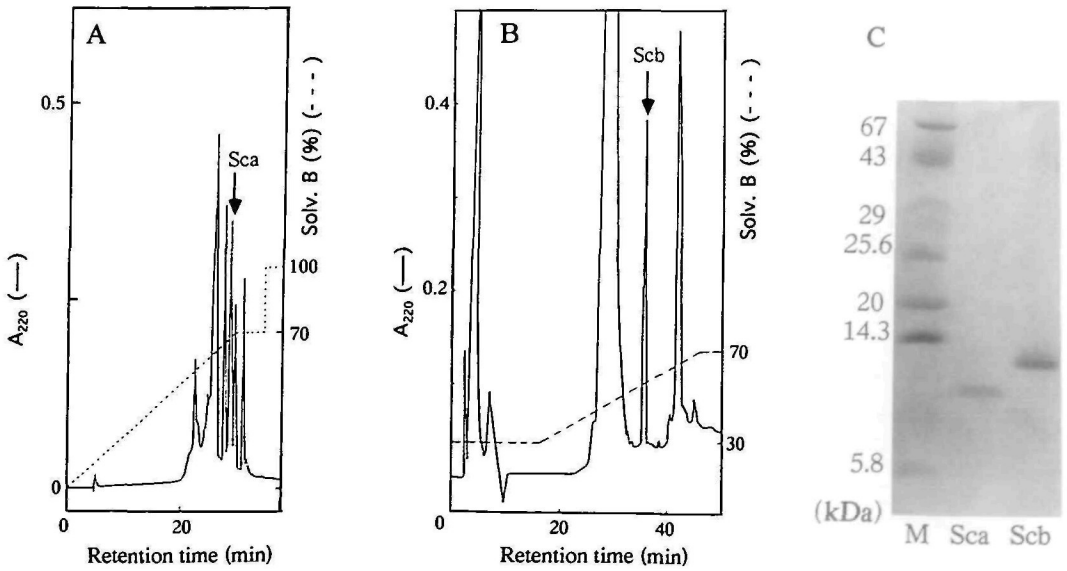

Fig. 4. **Purification of two sunflower cystatins, Sca and Scb, by RP-HPLC on a YMC GEL C4 column (4.6×250 mm).** The fractions exhibiting the inhibitory activities in Figs. 2B and 3 were lyophilized, dissolved in 0.1% TFA, and subjected to RP-HPLC. The proteins were eluted with increasing concentrations of the solvent B (80% acetonitrile in 0.1% TFA) at the flow rate of 1.0 ml/min. A and B indicate RP-HPLC patterns of Sca and Scb, respectively. C shows tricine SDS-PAGE of the purified Sca and Scb. The standard proteins used were BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), chymotrypsinogen (25.6 kDa), STI (20 kDa), lysozyme (14.3 kDa), and insulin (5.8 kDa).

ficin were estimated to be 1.9×10^{-6} and 2.8×10^{-9} M, respectively. This result showed that Scb has 30- and 700-fold higher affinities for papain and ficin, respectively, as compared to Sca.

Next, the effects of Sca and Scb on the activities of cathepsins B, H, and L purified from rats, as well as Arg-gingipain, an arginine-specific cysteine proteinase of *P. gingivalis*, were examined (Table II). Cathepsin H was strongly inhibited by Scb at a molar ratio of 1 : 1. Despite its weaker inhibitory potency, Sca exhibited significant inhibition of cathepsin H at an inhibitor/enzyme molar ratio of 10 : 1. Scb was significantly inhibitory for cathepsins B and L, whereas Sca had no inhibitory activity on these two enzymes. Arg-gingipain showed no inhibition by either Sca or Scb. In addition, neither Sca nor Scb inhibited bromelain, trypsin, or chymotrypsin (data not shown).

Amino Acid Sequence of Sca—Sca was first subjected to the N-terminal amino acid sequence analysis. This analysis determined the amino acid sequence up to the 24th residue, as shown in Fig. 6. Then, on the basis of the finding that Sca has one methionine residue, Sca was cleaved with CNBr and the resulting peptides were separated by RP-HPLC. As expected, two peptides CB1 and CB2 were isolated and they were directly sequenced (Fig. 6). Since sequence analyses of CB peptides were incomplete, Sca was cleaved by IBA at tryptophan residues, and two peptides (W1 and W2) were isolated and sequenced (Fig. 6). The sequence of peptide

TABLE II. **Effects of Sca and Scb on the activities of cathepsins B, L, H, Arg-gingipain, and papain.**

Inhibitor		Residual activity (%)[a]				
		Cathepsin B	Cathepsin L	Cathepsin H	Arg-gingipain	Papain
Sca	I/E=1.0	100	100	100	100	25.3
	=10.0	94.0	100	73.1	100	1.0
	=100.0	100	100	19.1	100	0.1
Scb	I/E=1.0	89.0	64.5	14.1	100	4.0
	=10.0	35.5	77.5	2.7	100	0.1
	=100.0	1.9	3.7	0.1	100	0.1

[a]Each purified enzyme was preincubated at 37°C for 10 min (cathepsins and papain) or 5 min (Arg-gingipain) with Sca and Scb at the indicated molar ratios, before 10 min incubation with the respective synthetic substrates (Z-Phe-Arg-MCA for cathepsins B and L, Arg-MCA for cathepsin H, Boc-Phe-Ser-Arg-MCA for Arg-gingipain, and Pyr-Phe-Leu-pNA for papain). The values are expressed as percentages of the activity determined in the absence of each inhibitor. The values determined for interaction with papain are also included for comparison.

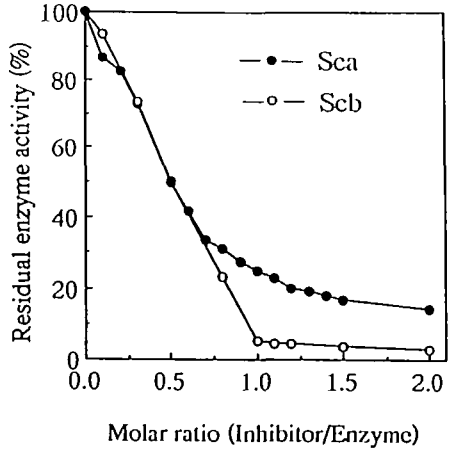

Fig. 5. **Inhibitory activities of Sca and Scb toward papain.** A fixed amount of enzyme was mixed with increasing amounts of the inhibitors, and the residual enzyme activities were assayed using Pyr-Phe-Leu-pNA as a substrate.

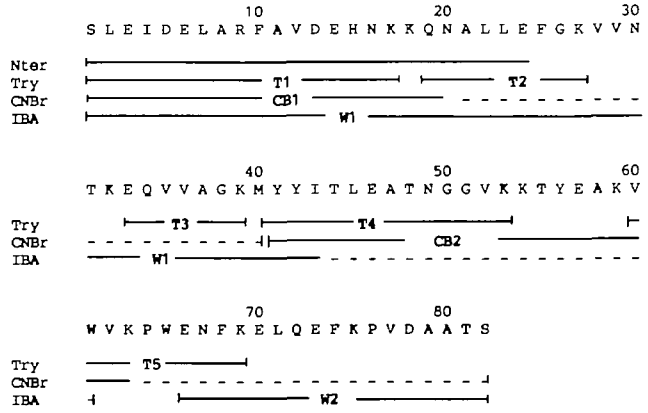

Fig. 6. **Amino acid sequence of the cysteine proteinase inhibitor Sca from sunflower seeds.** Try, CNBr, and IBA indicate peptides derived from tryptic, CNBr, and IBA cleavages, respectively. Nter indicates the amino acid residues determined by direct sequencing with a gas-phase sequencer, PSQ-1 (Shimadzu). —, indicates the amino acid residues identified by sequencing; and ····, shows the residues not identified by sequencing.

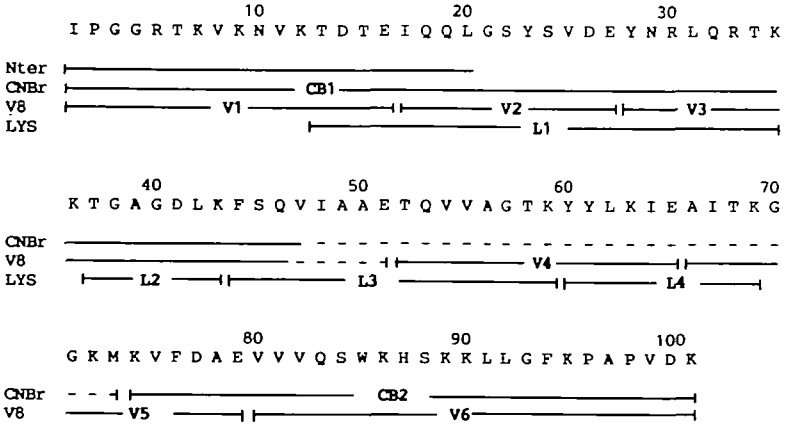

Fig. 7. **Amino acid sequence of the cysteine proteinase inhibitor Scb from sunflower seeds.** CNBr, V8, and LYS indicate the peptides derived from CNBr, *S. aureus* V8 proteinase, and lysylendopeptidase cleavages, respectively. Nter, —, and ···· indicate amino acid residues, as described in Fig. 6.

W1 could align the CB peptides, and information obtained from the CB and W peptides afforded almost the entire amino acid sequence of Sca. A few amino acids, however, remained undetermined. To establish the complete amino acid sequence of Sca, Sca was digested with trypsin and five peptides (T1-T5) were isolated by RP-HPLC, and sequenced. The amino acid sequence of the peptide T5 could align CB2 and W2, and in this way, the complete amino acid sequence of Sca was determined, as shown in Fig. 6. Sca consists of 83 amino acid residues and its molecular weight was calculated to be 9,330 Da.

Amino Acid Sequence of Scb—The amino acid sequence of Scb was established by procedures similar to those described above for Sca. The N-terminal sequence analysis of Scb using a gas phase sequencer could determine 20 amino acid residues (Fig. 7). Since Scb, like Sca, has one methionine residue, Scb was first cleaved with CNBr. The resulting peptides (CB1 and CB2) were separated by RP-HPLC and sequenced directly by a gas-phase sequencer (Fig. 7). This analysis gave the complete amino acid sequence of the C-terminal peptide CB2. To get sequence information of the middle region, Scb was digested with *S. aureus* V8 protease. The resulting peptides (V1-V6) were separated and sequenced (Fig. 7). Two peptides, V4 and V5, gave new sequence information; in particular, V5 contained the methionine residue, indicating that it might align the CB peptides. Since a few amino acids were still unidentified, Scb was further digested with lysylendopeptidase and four peptides (L1-L4), which were separated by RP-HPLC, were sequenced. The peptides L3 and L4 provided the overlaps of peptides V3-V4-V5, and could complete the amino acid sequence of Scb, as shown in Fig. 7. Scb consists of 101 amino acid residues and its molecular weight was calculated to be 11,187 Da from its sequence.

Sequence Comparison—The complete amino acid sequences of the two sunflower cystatins Sca and Scb are compared, as shown in Fig. 8. This comparison shows that Scb has 13 extra residues at the N-terminus, 7 insertions in the middle of the molecule, and lacks the C-terminal one residue when the residues of two proteins are aligned so as to optimize the similarity. The proteins Sca and Scb share 34 identical residues, that is, 42% of the residues compared. The two molecules have short runs of conserved sequences: Val^{25}-Asn^{29}, Gln^{53}-Tyr^{61}, and Phe^{94}-Pro^{96} in Scb, at the N-terminal, middle, and C-terminal regions, respectively.

Next, the amino acid sequences of Sca and Scb were aligned with those of various phytocystatins sequenced thus far (Fig. 9). In this comparison, Sca appears to lack about 10-25 amino acid residues at the N-terminus, as compared with other phytocystatins, even though it shows high homologies with other phytocystatins: oryzacystatin I (65%), oryzacystatin II (59%), avocado cystatin (72%), cowpea cystatin (73%), and corn cystatin (67%) (Fig. 9). These degrees of homology are comparable with those among other phytocystatins. By contrast, Scb shows relatively low similarities to phytocystatins: oryzacystatin I (28%), oryzacystatin II (33%), avocado cystatin (38%), cowpea cystatin (37%), and corn cystatin (37%). It also has a five amino acid insertion compared with Sca. Furthermore, the conserved residues among other phytocystatin molecules including Sca are often replaced with other amino acids in Scb, indicating a distinct structural character of Scb. Despite these alterations, there is a highly conserved sequence, Gln-Val-Val-Ala-Gly, in the middle region of all the molecules; this is thought to be one of the reactive sites for papain (*23*).

We further compared the sequences of Sca and Scb with

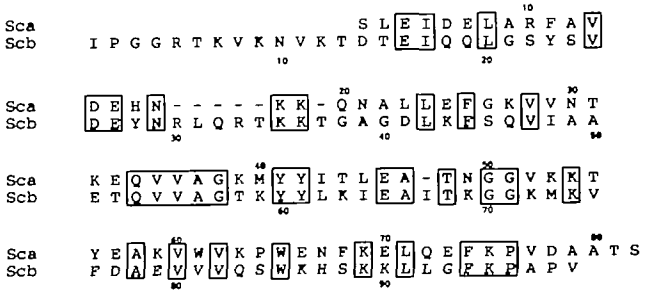

Fig. 8. **Sequence comparison of two sunflower cystatins Sca and Scb.** Identical residues are boxed.

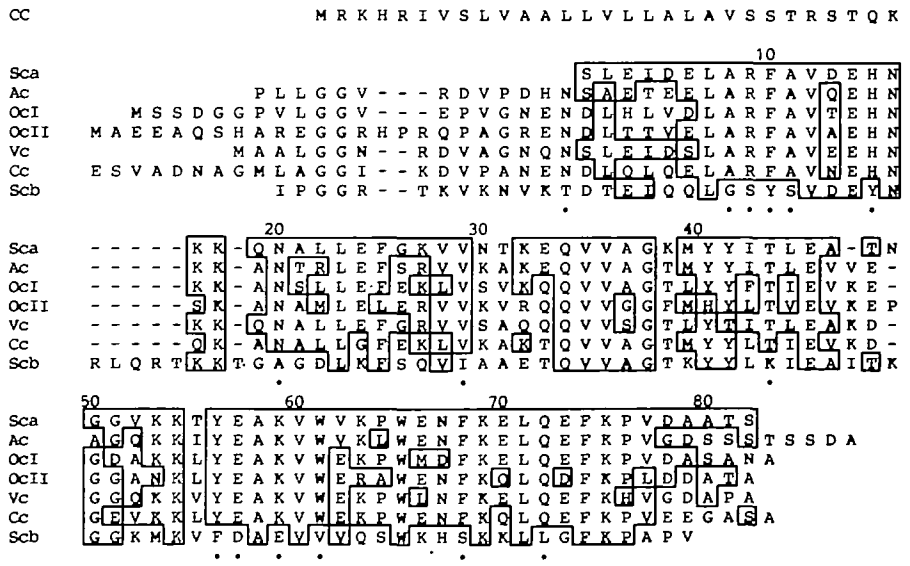

Fig. 9. **Sequence comparisons of the two sunflower cystatins with other phytocystatins.** The sequences of two sunflower cystatins were compared with those of oryzacystatins I (Oc I) and II (Oc II), avocado cystatin (Ac), cowpea cystatin (Vc), and corn cystatin (Cc). The residues identical to those in Sca are boxed. Dots indicate the conserved residues in all phytocystatins except for Scb.

those of animal cystatins, taking stefin B, chicken cystatin, and kininogen fragment as representative proteins of the three families. This comparison reveals that both sunflower cystatins, like other phytocystatins, show only low homologies to animal cystatins, sharing about 25% identical residues with chicken cystatin and kininogen fragment (families 2 and 3), and show even lower homology (20% identical residues) with stefin B (family 1) (data not shown). The result further shows that the degree of homology between Scb and animal cystatins is lower than those between Scb and phytocystatins. Thus, the sunflower cystatin Scb should be, in spite of its distinct structural features, a member of the phytocystatin family.

DISCUSSION

The protein structures of cysteine proteinase inhibitors from animal origin have been extensively investigated. These inhibitors have been classified into three families on the basis of sequence homology and the presence and position of intramolecular disulfide bonds: the stefins (family 1), which have M_r values of 11,000, lack disulfide bonds and carbohydrate; the cystatins (family 2), which have M_r values of 13,000, contain two intramolecular disulfide bonds, but lack any carbohydrate; and the kininogens (family 3), single-chain plasma proteins, whose N-terminal heavy chains are constructed of three glycosylated cystatin-like domains with inhibitory activity. In addition, Kondo *et al.* proposed, on the basis of sequence comparison of oryzacystatins I and II with animal cystatins, that plant cystatins should be classified into a fourth cystatin family "phytocystatin" (*5*). The structural properties common to phytocystatins were as follows: (1) the molecular weights are 10–12 kDa, (2) the disulfide bond is absent, and (3) the proteins are acidic or neutral proteins. The two cystatins, Sca and Scb, isolated from sunflower seeds have some characteristics different from other phytocystatins. First, Sca has a lower molecular weight (9.3 kDa) than other phytocystatins. Comparison of the sequence of Sca with those of other phytocystatins allowed assignment of the conserved Asn residue which precedes the N-terminal Ser residue in Sca (Fig. 9). Since an asparagine-specific post-translational processing enzyme occurs in plants (*24*), it is likely that the post-translation action of this processing enzyme between asparagine and serine residues would generate the mature Sca molecule, giving rise to the N-terminal sequence Ser-Leu-.

Recently, Bode *et al.* proposed the docking model in which papain interacted with cystatin forming the wedge composed by the residue Gly in the N-terminus as well as the amino acids Gln-Xaa-Val-Xaa-Gly in the middle of molecule and two amino acids Pro-Trp in the C-terminus (*23*). The inhibitory activity of rat cystatin α, lacking 6 residues from the N-terminus was one-eighth of that of the intact form and removal of the N-terminal 15 residues caused complete loss of the inhibitory activity (*25*). Similar observations concerning implication of the N-terminal region in the inhibitory activity were reported for both cystatin C and chicken cystatin (*26*). Furthermore, site-directed mutagenesis of cystatin A demonstrated the significance of the Gly residue at position 4, and it was concluded that the 4th residue must be small, like Gly, for interaction with the papain molecule (*27*). In contrast, it was reported that oryzacystatin I (*28*) and cystatin B (*29*) derivatives in which the N-terminal 21 and 6 amino acid residues, respectively, including the Gly residue, are missing, showed inhibitory activity toward papain. Hence, it was proposed, on the basis of these observations, that cystatins may be divided into two categories as to inhibitory activity; one group requires the N-terminal region and the other does not. Our sunflower cystatin Sca, which has no amino acid corresponding to Gly in the N-terminus, efficiently inhibits papain with a K_i value of 5.6×10^{-9} M, which is comparable to those determined for the interactions of oryzacystatins with papain (*5*). In this context, the sunflower cystatin Sca may be classified into the same category as oryzacystatin I and cystatin B.

An additional characteristic of the sunflower cystatins is that Scb is a highly basic protein with the pI value of 9.6, containing more basic residues (Lys_{13} and Arg_3) than other phytocystatins: other phytocystatins including Sca are acidic or neutral proteins, and the amino acid residues in Scb are often exchanged, compared to other phytocystatins. That is, the conserved amino acids in most phytocystatins, such as Ala^8, Arg^9, Phe^{10}, Ala^{11}, His^{12}, Asn^{17}, Leu^{20}, Glu^{21}, Val^{26}, Thr^{41}, Tyr^{53}, Glu^{54}, Lys^{56}, Trp^{58}, Phe^{65}, Gln^{68} in Sca, are replaced with different amino acids in Scb. In addition, six amino acid residues (Arg^{30}-Thr^{34}, Thr^{37}) are inserted in the middle of the Scb sequence as compared with other phytocystatins. The significance of these residues will be better understood when the three-dimensional structure of the phytocystatin molecule is known.

The inhibition spectra of phytocystatins against mammalian cysteine proteinases have been so far reported only for oryzacystatins (*5*), corn cystatin (*30*), and potato cysteine proteinase inhibitors (*9*). Oryzacystatins have the ability to inhibit only cathepsin H, while corn cystatin can inhibit cathepsins H and B with K_i values of 2.3×10^{-8} and 1.6×10^{-5} M, respectively. By contrast, the potato cysteine proteinase inhibitors strongly inhibited cathepsin L, and weakly inhibited cathepsins B and H. The inhibition profile of phytocystatin appears to be (from the highest affinity to the lowest) cathepsin H$\gg$cathepsin B$>$cathepsin L, while those of the potato cysteine proteinase inhibitors are cathepsin L$\gg$cathepsin B$>$cathepsin H, in general. In the present study, we investigated the interactions of the sunflower cystatins with the mammalian cysteine proteinases, cathepsins B, H, and L, as well as with plant cysteine proteinases, papain and ficin. The analysis of their inhibitory properties showed that although Sca inhibit plant cysteine proteinases, it has poor, if any, ability to inhibit the mammalian cysteine proteinases. In contrast to Sca, which has a narrow inhibitory spectrum against cysteine proteinases, Scb displays strong inhibition of the mammalian cysteine proteinases, and also inhibits plant cysteine proteinases, showing the inhibition profile of cathepsin H$\gg$cathepsin B$>$cathepsin L, which is similar to that of other phytocystatins.

Oryzacystatins I and II in the rice seeds are known as isocystatins. It has been reported that two oryzacystatins have complementary inhibitory activities toward papain and cathepsin H, although they are highly homologous with each other, showing a high sequence similarity (55% identity). Namely, oryzacystatin I strongly inhibits papain showing a K_i value of 3.0×10^{-8} M, while it weakly inhibits cathepsin H with a K_i value of 0.8×10^{-6} M, whereas

oryzacystatin II inhibits cathepsin H more strongly (1.0×10^{-8} M) than papain (0.8×10^{-6} M) (*5*). Thus, the two oryzacystatins are believed to be involved in the regulation of proteolysis caused by different proteinases. In addition, oryzacystatins I and II are thought to have a role as defense proteins by inhibiting cysteine proteinases derived from pathogens and/or insects. Sca and Scb in sunflower seeds are distinct in protein structure and inhibitory spectrum. The inhibitory properties displayed by Sca and Scb suggest that Sca may have a function in the sunflower seed as a regulator of an endogenous papain-like enzyme, and that Scb might be involved in protection against exogenous enzymes resembling cathepsins produced by insects and pathogens. Whichever functions they have, it seems likely that the difference in affinity for cysteine proteinases of Sca and Scb found in sunflower seeds is the result of an evolutionary process, broadening the total inhibitory capacity in the seeds.

REFERENCES

1. Barrett, A.J., Rawlings, N.D., Davies, M.E., Machleidt, W., Salvesen, G., and Turk, V. (1986) Cysteine proteinase inhibitors of the cystatin superfamily in *Research Monographs in Cell and Tissue Physiology*, Vol. 12, *Proteinase Inhibitors* (Barrett, A.J. and Salvensen, G., eds.) pp. 515-570, Elsevier, Amsterdam, New York, Oxford
2. Abe, K. and Arai, S. (1985) Purification of a cysteine proteinase inhibitor from rice, *Oryza sativa* L. japonica. *Agric. Biol. Chem.* **49**, 3349-3350
3. Abe, K., Kondo, H., and Arai, S. (1987) Purification and characterization of a rice cysteine proteinase inhibitor. *Agric. Biol. Chem.* **51**, 2763-2768
4. Abe, K., Emori, Y., Kondo, H., Suzuki, K., and Arai, S. (1987) Molecular cloning of a cysteine proteinase inhibitor of rice (Oryzacystatin). *J. Biol. Chem.* **262**, 16793-16797
5. Kondo, H., Abe, K., Nishimura, I., Watanabe, H., Emori, Y., and Arai, S. (1990) Two distinct cystatin species in rice seeds with different specificities against cysteine proteinases. *J. Biol. Chem.* **265**, 15832-15837
6. Fernandes, K.V.S., Sabelli, P.A., Barratt, D.H.P., Richardson, M., Xavier-Filho, J., and Shewry, P.R. (1993) The resistance of cowpea seeds to bruchid beetles is not related to levels of cysteine proteinase inhibitors. *Plant Mol. Biol.* **23**, 215-219
7. Krizaj, I., Drobnic-Kosorok, M., Brzin, J., Jerala, R., and Turk, V. (1993) The primary structure of inhibitor of cysteine proteinases from potato. *FEBS Lett.* **333**, 15-20
8. Kimura, M., Ikeda, T., Fukumoto, D., Yamasaki, N., and Yonekura, M. (1995) Primary structure of a cysteine proteinase inhibitor from the fruit of avocado (*Persea americana* Mill). *Biosci. Biotech. Biochem.* **59**, 2328-2329
9. Brzin, J., Popovic, T., Drobnic-Kosorok, M., Kotnik, M., and Turk, V. (1988) Inhibitors of cysteine proteinases from potato. *Biol. Chem. Hoppe-Seyler* **369**, Suppl., 233-238
10. Abe, M., Abe, K., Kuroda, M., and Arai, S. (1992) Corn kernel cysteine proteinase inhibitor as a novel cystatin superfamily member of plant origin. *Eur. J. Biochem.* **209**, 933-937
11. Yamamoto, K., Kamata, O., and Kato, Y. (1984) Separation and properties of three forms of cathepsin H-like cysteine proteinase from rat spleen. *J. Biochem.* **95**, 477-484
12. Yamamoto, K., Kamata, O., Takeda, M., and Kato, Y. (1983) Characterization of rat spleen cathepsin B and different effects of anti-inflammatory drugs on its activity. *Jpn. J. Oral Biol.* **25**, 834-838
13. Yamamoto, K., Takeda, M., and Kato, Y. (1985) Characteristics of activation of cathepsin B by sodium salicylate and comparison of catalytic site properties of cathepsin B and H. *Jpn. J. Pharmacol.* **39**, 207-215
14. Bando, Y., Kominami, E., and Katunuma, N. (1986) Purification and tissue distribution of rat cathepsin L. *J. Biochem.* **100**, 35-42
15. Kadowaki, T., Yoneda, M., Okamoto, K., Maeda, K., and Yamamoto, K. (1994) Purification and characterization of a novel arginine-specific cysteine proteinase (argingipain) involved in the pathogenesis of periodental disease from the culture supernatant of *Porphyromonas gingivalis*. *J. Biol. Chem.* **269**, 21371-21378
16. Schagger, H. and Jagow, G.V. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379
17. Svensson, H. (1962) Isoelectric fractionation, analysis, and characterization of ampholytes in natural pH gradients. *Acta Chem. Scand.* **16**, 456-466
18. Filippova, I.Y., Lysogorskaya, E.N., Oksenoit, E.S., Rudenskaya, G.N., and Stephanov, V.M. (1984) L-Pyroglutamyl-L-phenylalanyl-L-leucine-*p*-nitroanilide, a chromogenic substrate for thiol proteinase assay. *Anal. Biochem.* **143**, 293-297
19. Steers, E., Craven, G.R., Jr., Anfinsen, C.B., and Bethune, J.L. (1965) Evidence for nonidentical chains in the β-galactosidase of *Escherichia coli* K12. *J. Biol. Chem.* **240**, 2478-2484
20. Fontana, A., Dalzoppo, D., Grandi, C., and Zambonin, M. (1983) Cleavage at tryptophan with *o*-iodosobenzoic acid. *Methods Enzymol.* **91**, 311-318
21. Heinrikson, R.L. and Meredith, S.C. (1984) Amino acid analysis by reverse-phase high performance liquid chromatography: Precolumn derivatization with phenylisothiocyanate. *Anal. Biochem.* **136**, 65-74
22. Crestifield, A.M., Moore, S., and Stein, W.H. (1963) The preparation and enzymatic hydrolysis of reduced and *S*-carboxymethylated proteins. *J. Biol. Chem.* **238**, 622-627
23. Bode, W., Enbh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **7**, 2593-2599
24. Hara-Nishimura, I., Shimada, T., Hiraiwa, N., and Nishimura, M. (1995) Vacuolar processing enzyme responsible for maturation of seed proteins. *J. Plant Physiol.* **145**, 632-640
25. Takeda, A., Kaji, H., Nakaya, K., Aoki, Y., Nakamura, Y., and Samejima, T. (1985) Amino acid sequence of derivatives of newborn rat epidermal thiol proteinase inhibitor. *Biochem. Int.* **11**, 557-564
26. Abrahamson, M., Ritonja, A., Brown, M.A., Grubb, A., Machleidt, W., and Barrett, A.J. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors, human cystatin C and chicken cystatin. *J. Biol. Chem.* **262**, 9688-9694
27. Shibuya, K., Kaji, H., Ohyama, Y., Tate, S., Kainosho, M., Inagaki, F., and Samejima, T. (1995) Significance of the highly conserved Gly-4 residue in human cystatin A. *J. Biochem.* **118**, 635-642
28. Abe, K., Emori, Y., Kondo, H., Arai, S., and Suzuki, K. (1988) The NH_2-terminal 21 amino acid residues are not essential for the papain-inhibitory activity of oryzacystatin, a member of the cystatin superfamily. *J. Biol. Chem.* **263**, 7655-7659
29. Thiele, U., Assfalg-Machleidt, I., Machleidt, W., and Auerswald, E.A. (1990) N-terminal variants of recombinant stefin B: Effect on affinity for papain and cathepsin B. *Biol. Chem. Hoppe-Seyler* **371**, Suppl., 125-136
30. Abe, M. and Arai, S. (1991) Some properties of a cysteine proteinase inhibitor from corn endosperm. *Agric. Biol. Chem.* **55**, 2417-2418