J. Biochem. 124, 911-916 (1998)

Molecular Cloning, Functional Expression, and Mutagenesis of cDNA Encoding a Cysteine Proteinase Inhibitor from Sunflower Seeds

Keiko Doi-Kawano, Yoshiaki Kouzuma,[1] Nobuyuki Yamasaki, and Makoto Kimura

Laboratory of Biochemistry, Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581

Received for publication, May 25, 1998

Sunflower cystatin Scb differs from other phytocystatins in that it is a highly basic protein with a pI value of 9.6 and includes six additional amino acids (Arg^{30}-Leu-Gln-Arg-Thr^{34}, Thr^{37}) in the middle region as compared with other phytocystatins [Kouzuma *et al.* (1996) *J. Biochem.* 119, 1106-1113]. We identified and sequenced a complete cDNA encoding the Scb; the cDNA of Scb consists of 645 nucleotides and includes an open reading frame encoding a polypeptide of 123 amino acids. On the basis of these findings, Scb appears to be synthesized as a prepeptide consisting of a signal sequence of 22 amino acids and a mature protein of 101 amino acids. A recombinant Scb (rScb) was produced by expression in *Escherichia coli* and purified by gel filtration on Sephacryl S-200 followed by ion-exchange column chromatography on a S-Sepharose column. rScb exhibited almost the same inhibitory activity toward papain as the authentic Scb did, but its inhibition profile toward cathepsins B, L, and H was slightly different. Scb mutant proteins, in which selected N-terminal residues or the additional amino acids were deleted, were subsequently constructed and characterized with respect to their inhibitory activities toward papain. The result revealed that the additional sequence (Arg^{30}-Leu-Gln-Arg-Thr^{34}) in Scb is not essential for papain-inhibitory activity, while the N-terminal amino acids (Ile^{1}-Pro^{2}) as well as the N-terminal glycine residues Gly^{3} and/or Gly^{4} play an important role in manifesting the inhibitory activity toward papain.

Key words: cDNA cloning, cysteine proteinase, overexpression, phytocystatin, sunflower.

Since oryzacystatin I, which is the best-characterized cysteine proteinase inhibitor from plants, was isolated from rice seeds (*1*), some sequence information on cystatins of plant origin, the so-called phytocystatins, has become available either by deduction from cDNA (*2-6*) sequences or by direct protein sequencing (*7*). The following structural properties are common to phytocystatins. (i) Their molecular weights are 10-12 kDa, (ii) the disulfide bond is absent, (iii) the proteins are either acidic or neutral, and (iv) like animal cystatins, they share three conservative motifs: Gly in the vicinity of the N-terminal region, Gln-Xaa-Val-Xaa-Gly in the middle of the molecule, and Pro-Trp in the C-terminal region.

We previously isolated two distinct phytocystatins, designated Sca and Scb, from sunflower seeds and characterized their primary structures and inhibitory specificities (*8*). Both sunflower cystatins have structural features in common with phytocystatins, including a low molecular weight, the absence of a cysteine residue, and strict conservations of Q-X-V-X-G and PW residues. However, the study also revealed that Scb differs from Sca as well as from other phytocystatins in that it is a highly basic protein with a pI value of 9.6 and that it has six additional amino acids in the middle of the molecule. Furthermore, unlike Sca, Scb has an N-terminal extension composed of 13 amino acids.

In the study reported here, in order to elucidate the roles of the N-terminal and the additional Scb residues in its inhibitory activity, we cloned a cDNA encoding Scb and overproduced it in a functional form in *Escherichia coli* cells. Subsequently, several Scb mutants, in which N-terminal amino acids or the additionl sequence were deleted, were constructed and their inhibitory activities toward papain were investigated.

MATERIALS AND METHODS

Materials—Sunflowers were grown on the experimental farm of Kyushu University and developing sunflower seeds were harvested at 10-day intervals after flowering. The seeds were frozen in liquid nitrogen and stored at −80°C until used. The oligonucleotides used in this study were purchased from Greiner Japan. A predigested λgt11/*Eco*RI vector kit and a Time Saver cDNA synthesis kit were obtained from Pharmacia. A Gigapack™ II Plus phage

[1] To whom correspondence should be addressed. Tel: +81-92-642-2854, Fax: +81-92-642-2854, E-mail: kouzuma@agr.kyushu-u.ac.jp

Abbreviations: IPTG, isopropyl-β-D-thiogalactopyranoside; MALDI-TOF, a matrix assisted laser desorption and ionization time of flight; PCR, polymerase chain reaction; rScb, recombinant Scb; RP-HPLC, reverse-phase high-performance liquid chromatography; RT-PCR, reverse transcription-PCR; Sca, sunflower cystatin a; Scb, sunflower cystatin b; SDS-PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid.

packaging kit was purchased from Stratagene. *Taq* DNA polymerase was obtained from Sawady Technology. A random primer DNA labeling kit, DNA ligation kit, and Takara Taq cycle sequencing kit were obtained from Takara Shuzo and used as recommended by the supplier. Restriction endonucleases and DNA modifying enzymes were purchased from either Toyobo, MBI Fermentas, or Gibco BRL. The plasmid vectors used in this work were as follows: pUC18 from MBI Fermentas, pGEM™-T vector from Promega, and pET-22b expression vector from Novagen. Nitorocellulose membranes were from Schleicher & Schuell. [α-^{32}P]dCTP was obtained from Amersham. Papain was obtained from Sigma. All other chemicals were of analytical grade for biochemical use.

Extraction of Total RNA and Construction of a cDNA Library—Total RNA was extracted from the developing sunflower seeds by the method of Manning (*9*). Poly(A)$^+$ RNA was purified by utilizing Oligotex™-dT30 〈Super〉 (Takara Shuzo). Subsequently, a cDNA library was constructed with the cDNA synthesis kit and λ gt11 as a cloning vector according to the manufacture's protocol.

Preparation of a Scb Gene Fragment by PCR—The two oligonucleotides with an *Eco*RI site were synthesized for use as primers corresponding to the amino acid sequences of Scb: 5′-GTNAARACNGAYACNGAATTCCARCA-3′ (VKTDTEIQQ at positions 11-19) and 5′-GARGTNGTN-GTNGAATTCTGGAARCA-3′ (EVVVQSWKH at positions 79-87), where R represents A and G; Y, C and T; N, A, C, G, and T. RT-PCR was performed with developing sunflower seed cDNAs as a template on an automated thermal cycler (Astec Program Temp. Control System PC-700). The PCR products were purified on a gel, treated with *Eco*RI, and then subcloned into pUC18, previously digested with *Eco*RI. After confirmation of its sequence, the fragment was used as a hybridization probe to screen the cDNA library.

Screening and Nucleotide Sequencing—The developing sunflower seed cDNA library was screened by plaque hybridization using the PCR product labeled by the random priming method, as described in Ref. *10*. By the second and third plaque hybridizations, positive clones were isolated and phage DNAs were prepared from them. The cDNA inserts were excised by *Eco*RI, subcloned into the plasmid vector pUC18, and sequenced with a DNA sequencer DSQ-1000 (Shimadzu) using the Takara Taq cycle sequencing kit.

Expression of Scb in E. coli Cells—To obtain a cDNA fragment encoding a mature Scb, PCR was performed using the Scb cDNA as a template with the forward primer 5′-CATATGATTCCAGGAGGAAGAACTAAG-3′ and reverse primer 5′-AATGGATCCTTACTTATCAACCGGT-GCCGG-3′. The PCR products were ligated into the pGEM™-T vector. After confirmation of the DNA sequence, the DNA fragment encoding the mature Scb was excised by digestion with *Nde*I and *Bam*HI, and ligated into the expression vector pET-22b, previously digested with the same enzymes. The resulting plasmid, pET-Scb, was introduced into *E. coli* BL21(DE3) and recombinant protein was induced with 1 mM IPTG according to the supplier's instruction.

SDS-Polyacrylamide Gel Electrophoresis and Immunoblotting—Overproduction of Scb in *E. coli* cells was analyzed by SDS-PAGE using 15% polyacrylamide gels (*11*) and Western blotting (*12*). In the Western blotting analysis, the nitorocellulose filter was probed with polyclonal anti-Scb antibodies, developed using goat anti-mouse IgG conjugated with horseradish peroxidase, and the immune complex was then visualized by addition of a peroxidase substrate solution, 100 mM Tris-HCl, pH 7.5, containing 0.5 mg/ml 3,3′-diaminobenzidine tetrahydrochloride, 0.3% H_2O_2, and 50 mM imidazole.

Purification of the Recombinant Scb (rScb)—*E. coli* transformants harbouring the expression vector pET-Scb were cultured in 1 liter of LB-broth with 50 μg/ml ampicillin. After cultivation, cells were harvested by centrifugation and washed twice in 30 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl and resuspended in the same buffer. Cells were sonicated on ice at a setting of 7 for 2-min periods followed by 5 min for cooling for a total of 70 min. The soluble and insoluble fractions were separated by centrifugation at 12,000 × *g* for 15 min at 4°C.

The soluble fraction was salted out by saturation with ammonium sulfate. The precipitate was recovered, dissolved in 30 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl and 2 M urea, and dialyzed against the same buffer. The protein solution was chromatographed on Sephacryl S-200 (2.8 × 80 cm) equilibrated with the same buffer. Each fraction was subjected to SDS-PAGE; the fractions showing the same molecular mass as Scb were collected and further purified on S-Sepharose (1.5 × 15 cm) equilibrated with 10 mM Na-phosphate buffer (pH 7.2). The proteins were eluted with a linear gradient of 0 to 0.5 M NaCl in the same buffer. The fractions showing papain inhibitory activity were collected and lyophilized. Proteins were dissolved in 10 mM Na-phosphate buffer, pH 7.2, and finally purified by RP-HPLC on a Wakosil 5C_{18} (300 Å) column (4.6 × 250 mm) with 0.1% TFA. The protein was eluted with a linear gradient of 24 to 56% acetonitrile in 0.1% TFA.

The insoluble fraction was washed with 1 M sucrose followed by 2% Triton X-100/10 mM EDTA. The inclusion body thus obtained was dissolved in 30 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl and 7 M guanidine hydrochloride. The protein solution was dialyzed against 30 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl and 2 M urea, and the protein solution was purified by the same procedures as those used for the protein in the soluble fraction. The molecular mass of the recombinant was measured by MALDI-TOF mass spectrometer Voyager RP (PerSeptive Biosystems). The N-terminal amino acid sequence of the protein was analyzed by a gas phase sequencer PSQ-1 (Shimadzu).

Proteinase Inhibition Assays—Papain, and cathepsins B,

TABLE I. **Scb mutants and the sequences of their oligonucleotide primers.**

Mutant	Oligonucleotide primer
rScb	5′-CATATGATTCCAGGAGGAAGAACTAAG-3′
	M I P G G R T K
Δ5	5′-CAGTGGACGAGTACAATAAGAAAACCGGCGCCGG-3′
	S V D E Y N K K T G A G
N2	5′-CATATGGGAGGAAGAACTAAGGTCAAAAACGTC-3′
	M G G R T K V K N V
N4	5′-CATATGAGAACTAAGGTCAAAAACGTCAAAACCG-3′
	M R T K V K N V K T
N13	5′-CATATGGACACTGAGATTCAGCAACTCGG-3′
	M D T E I Q Q L D

L, and H inhibition assays were performed as described in previous by Ref. *8*. Concentrations of Scb and the mutant proteins were determined by amino acid analysis.

Construction of Scb Mutants—Site-directed mutagenesis was performed by the unique site elimination method developed by Deng and Nickoloff (*13*) using a Chameleon™ double-stranded site-directed mutagenesis kit (Stratagene). The N-terminal deletion mutants were constructed by the PCR method with specific primers. The nomenclature of the Scb mutants and the oligonucleotide primers used for mutagenesis are shown in Table I. Mutations by the unique site elimination method were introduced into the amplified Scb cDNA fragment that had been subcloned into the pGEM™-T vector. The N-terminal deletion mutants constructed by the PCR method were subcloned into the pGEM™-T vector. After mutagenesis, the Scb cDNA fragments were sequenced to verify the presence of the desired mutation. The mutant fragments were then recovered and ligated to the expression vector pET-22b. Expression and purification of all the mutants were done by procedures identical to those described for rScb.

RESULTS

Cloning and Nucleotide Sequence of the Scb cDNA—Developing sunflower seeds were harvested, mRNA was extracted, and the corresponding cDNA was synthesized. With the help of two degenerate primers as described in "MATERIALS AND METHODS," one fragment (230 bp) was reproducibly amplified by PCR (data not shown). This was the expected size based on the positions of the primers in the protein sequence. After confirmation of its nucleotide sequence, this fragment was used to screen the Scb cDNA clone. Screening of approximately 1×10^5 independent recombinant clones from a λgt11 cDNA library gave six positive clones whose insert cDNA fragments were subcloned into the pUC18 vector and sequenced.

The nucleotide sequence of the cDNA clone containing the largest insert (645 bp) and its deduced amino acid sequence are shown in Fig. 1. The cDNA insert is composed of a 129-bp 5′ noncoding sequence, a 369-bp coding sequence for Scb, a stop codon (TAA), and a 144-bp 3′ noncoding sequence. Comparison of the deduced amino acid sequence from the nucleotide sequence of Scb cDNA with the protein sequence reported previously (*8*) revealed that

```
  1  CCG TGT CCT CTT TAT CCA AAA ACC CGC ATC AAA TTC GGT GCC TCT TGT TTT CGC TTT TCA   60

 61  ACA AGC TTC TTT CCC CTA TAC ATA TAC ACA CAC CCA CTT CAT CTT CAA TCA AAC AAA CCA  120

121  AAT TTC ACC ATG TCC TCC AAA CTC CCA ATA ACC TTC TTC ATC ATT TCG CTC TCC CTT TTC  180
-22            Met Ser Ser Lys Leu Pro Ile Thr Phe Phe Ile Ile Ser Leu Ser Leu Phe   -6

172  CTC GTG GCA ACC GCC ATT CCA GGA GGA AGA ACT AAG GTC AAA AAC GTC AAA ACC GAC ACT  240
 -5  Leu Val Ala Thr Ala Ile Pro Gly Gly Arg Thr Lys Val Lys Asn Val Lys Thr Asp Thr   15

241  GAG ATT CAG CAA CTC GGA AGT TAC TCA GTG GAC GAG TAC AAT CGG TTA CAA CGG ACG AAG  300
 16  Glu Ile Gln Gln Leu Gly Ser Tyr Ser Val Asp Glu Tyr Asn Arg Leu Gln Arg Thr Lys   35

301  AAA ACC GGC GCC GGA GAT CTG AAA TTC TCG CAG GTG ATT GCG GCG GAG ACG CAG GTG GTG  360
 36  Lys Thr Gly Ala Gly Asp Leu Lys Phe Ser Gln Val Ile Ala Ala Glu Thr Gln Val Val   55

361  GCC GGA ACC AAA TAT TAC TTG AAG ATT GAA GCG ATT ACG AAG GGC GGT AAG ATG AAG GTG  420
 56  Ala Gly Thr Lys Tyr Tyr Leu Lys Ile Glu Ala Ile Thr Lys Gly Gly Lys Met Lys Val   75

421  TTC GAT GCG GAA GTG GTG GTT CAG TCA TGG AAG CAT TCG AAG AAG TTG TTA GGG TTT AAA  480
 76  Phe Asp Ala Glu Val Val Val Gln Ser Trp Lys His Ser Lys Lys Leu Leu Gly Phe Lys   95

481  CCG GCA CCG GTT GAT AAG TAA ATA TTA ATT TTG TGT TTT GTG TGT GTG TGT TGA GAT CTG  540
 96  Pro Ala Pro Val Asp Lys TER                                                    101

541  TTT GGT TTA TGG TGT GGT AAC TCG GTT TTT TTT TAT GAA TAA TAT GTA AGG ATC TGT GGC  600

601  TGA TTA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA                  645
```

Fig. 1. **Nucleotide and deduced amino acid sequence of Scb cDNA clone.** The predicted amino acid sequence is shown below the nucletide sequence. The arrowhead indicates the processing site of Scb. The signal peptide at the N-terminal is boxed.

the deduced sequence has a putative signal sequence consisting of 22 amino acid residues, and that the amino acid sequence of the deduced mature protein is completely identical with that determined by the protein sequencing.

Expression of Scb in E. coli Cells—The expression plasmid pET-Scb was introduced into *E. coli* strain BL21(DE3) and the expression of Scb was induced by adding IPTG to the culture medium. Western blotting analysis of the cell lysate of *E. coli* revealed a 14-kDa protein band in the *E. coli* extracts induced with IPTG, but not in the extracts without induction (data not shown). Subsequently, the extract from *E. coli* expressing Scb was prepared by sonication, and the soluble and insoluble fractions were analysed by SDS-PAGE followed by Western blotting. This analysis showed that the recombinant Scb (rScb) occurred in both the cytoplasmic fraction and inclusion bodies.

Purification of the Recombinant Scb (rScb)—In preliminary experiments, when the *E. coli* lysate induced with IPTG was put directly onto Sephacryl S-200, the rScb was eluted at void volumes, suggesting that the rScb obtained in the cytoplasmic fraction was oligomerized, possibly through hydrophobic interaction. Thus, in the subsequent experiments, the rScb in the soluble fraction was purified in the presence of 2 M urea in the solution. The soluble fraction was prepared after sonication of *E. coli* cells and rScb was denatured in a buffer containing 2 M urea, as described in "MATERIALS AND METHODS." The rScb was purified by gel filtration on Sephacryl S-200 followed by ion-exchange column chromatography on a S-Sepharose column. The inhibitor fraction was finally purified by RP-HPLC on a Wakosil $5C_{18}$ column (Fig. 2). The major peak showed strong papain-inhibitory activity and had almost the same molecular mass (11,219 Da) measured by a MALDI-TOF mass spectrometer. The yield of rScb was about 10 mg/liter of induced culture.

The rScb expressed in inclusion bodies was dissolved in 7 M guanidine hydrochloride and refolded by dialyzation against a buffer containing 2 M urea. Purification of the rScb from inclusion bodies was performed as described under "MATERIALS AND METHODS," and the protein was obtained to the same amount as that from the soluble fraction.

Sequence analysis of the recombinant proteins isolated from the soluble fraction and inclusion bodies gave identical N-terminal sequences: Met-Ile-Pro-Gly-, indicating that both rScb proteins have an extra Met residue attached to the N-terminus of Scb.

The inhibitory potency of the rScb isolated from both the soluble fraction and inclusion bodies toward the cysteine proteinases papain and cathepsins were examined. The result showed that the rScb from both fractions inhibited papain to almost the same extent as the authentic Scb, suggesting that both rScb proteins are folded into an active conformation. Thus, in further studies, the rScb isolated from the soluble fraction were used as a wild-type Scb. Subsequently, the inhibitory activity of rScb toward cathepsins was examined. The rScb exhibited slightly different inhibitory activity from the authentic Scb toward cathepsins: it inhibited cathepsin B more strongly than the authentic Scb did, while it showed weaker inhibitory activity toward cathepsins L and H (Table II). It can be postulated that the difference in inhibitory activities toward cathepsins shown by rScb and Scb might be attributable to the extra methionine residue at the N-terminus of the rScb.

Construction and Characterization of Scb Mutants—To investigate the structure-activity relationship of Scb, four Scb mutant cDNAs were prepared by *in vitro* mutagenesis and expressed in *E. coli* cells. The resulting mutant proteins were purified from the cytoplasmic fraction by procedures identical to those described for rScb. Figure 3 shows SDS-PAGE profiles of typical preparations of rScb and its mutant proteins. Three deletion mutants, N2, N4, and N13, in which the N-terminal was truncated by two, four, and thirteen amino acids, respectively, were constructed to examine the significance of the N-terminal residues in Scb for papain inhibitory activity. A fourth

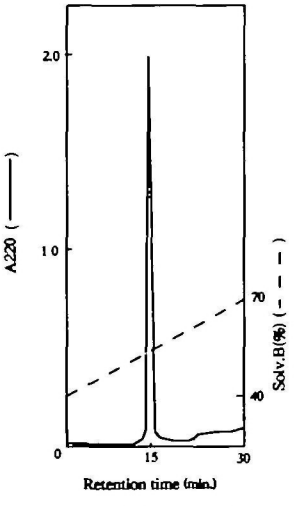

Fig. 2. **Reverse-phase HPLC of purified recombinant Scb.** Proteins obtaind by S-Sepharose column chromatography were subjected to RP-HPLC on a Wakosil $5C_{18}$ (300 Å) column (4.6×250 mm) with 0.1% TFA at a flow rate of 1 ml/min. The proteins were eluted with a linear gradient of 24 to 56% acetonitrile in 0.1% TFA.

TABLE II. **Inhibitory activity of authentic and recombinant Scbs toward cysteine proteinases.**

	Residual activity (%)			
	Papain	Cathepsin B	L	H
Authentic Scb	7.2	35.5	40.3	3.4
Recombinant Scb	7.3	26.2	50.7	37.3

Inhibitor/Papain = 1.0. Inhibitor/Cathepsins = 10.0.

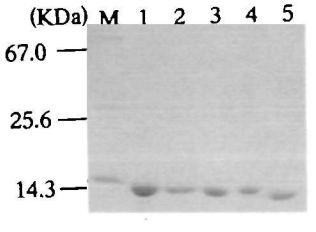

Fig. 3. **SDS-polyacrylamide gel electrophoresis of purified rScb and its mutant.** Samples were applied as follows: rScb (lane 1), N2 (lane 2), N4 (lane 3), N13 (lane 4), and Δ5 (lane 5). M indicates the standard protein, which were bovine serum albumin (67.0 kDa), α-chymotrypsinogen (25.6 kDa), and lysozyme (14.3 kDa).

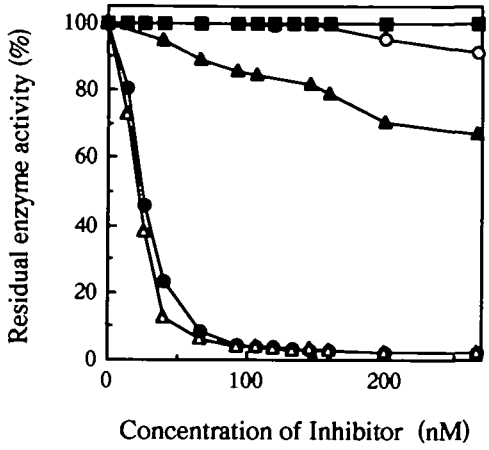

Fig. 4. **Inhibition of papain by increasing amounts of rScb and its mutant proteins.** A fixed amount of papain was mixed with increasing amounts of the inhibitors, and the residual enzyme activities were assayed as described in Ref. *8*. Symbols: rScb (●); Δ5 (△); N2 (▲); N4 (○); and N13 (■).

mutant was also constructed (Δ5) in which five of the six amino acids in the additional sequence in Scb (Arg^{30}-Leu-Gln-Arg-Thr^{34}) was deleted to investigate its importance for the inhibitory activity toward papain.

As shown in Fig. 4, the inhibitory activity of the mutant N2 was significantly decreased as compared with that of rScb, while N4 and N13 scarcely retained any inhibitory activity toward papain. These results suggest that the glycine residues (Gly^3 and/or Gly^4) and the preceding residues (Ile^1-Pro^2), in the N-terminal region are very important for papain inhibition. In contrast, the mutant designated Δ5, in which the additional sequence (Arg^{30}-Leu-Gln-Arg-Thr^{34}) was deleted, retained full inhibitory activity toward papain. This indicates that the characteristic additional amino acid residues in Scb are not involved in inhibitory activity toward papain.

DISCUSSION

We isolated a cDNA clone encoding the phytocystatin Scb from sunflower seeds. The Scb cDNA comprises 645 nucleotides and the open reading frame encodes a polypeptide of 123 amino acids. This includes a possible signal peptide composed of 22 amino acids which resembles a classical signal sequence: it has charged residues very near the N-terminus followed by a long stretch of hydrophobic amino acids and ends with an amino acid possessing a small side chain (*14*). Although the localization of Scb in cells is still not known, this finding suggests that Scb may be synthesized as a pre-Scb consisting of a signal sequence of 22 amino acid residues which must be proteolytically processed at its functional site.

The cDNA encoding the mature Scb was expressed in *E. coli* cells. rScb was detected in both the supernatant and the inclusion bodies, and the rScb proteins obtained from both fractions, each of which had an extra Met at the N-terminus, could inhibit papain to the same extent as the authentic Scb. This indicates that rScb could refold into a conformation similar to that of the authentic Scb, and also that the additional methionine in the N-terminus had no effect on the papain-inhibitory activity of rScb. This result is well consistent with the fact that there is no difference between the circular dichroism spectra of the authentic Scb and rScb (data not shown). On the other hand, the inhibitory activities of rScb toward cathepsins from mice were slightly different. rScb showed stronger inhibitory activity toward cathepsin B, and weaker activity toward cathepsins L and H than the authentic Scb did. Cathepsins B, L, and H were reported to be inhibited by cystatin C to different extents, and the K_i values toward these cathepsins were increased or decreased by mutagenesis of the N-terminal region of cystatin C (*28*). It was suggested that the different effects on cathepsins might be attributable to their ability to recognize the N-terminal region of cystatin. Therefore, it is likely that the different inhibitory activities toward cathepsins shown by rScb compared to those of Scb might be due to the extra methionine residue attached at the N-terminus of rScb.

A crystallographic study of chicken cystatin suggested that the proteinase binding site comprises three regions, the N-terminal region around Gly^9, the Gln-Leu-Val-Ser-Gly sequence at residues 53-57, and the region around Trp^{104} (*15*). This proposition was confirmed by a crystallographic study of a complex consisting of cystatin B and papain, in which cystatin B interacts with papain to form a wedge composed of three motifs corresponding to those found in the structure of chicken cystatin (*16*). However, a cystatin B derivative, in which 6 N-terminal amino acids including the Gly residue were missing, was found to exhibit inhibitory activity (*17*), indicating that the N-terminal region is dispensable for inhibition. In contrast, implication of the N-terminal region in the inhibitory activity was reported for cystatin C (*18-20*) and chicken cystatin (*21*). Furthermore, site-directed mutagenesis of cystatin A demonstrated the significance of the Gly residue at position 4 (*22, 23*). Meanwhile, Shibuya *et al.* found by NMR spectroscopic analysis of cystatin A that the mutation of the Gly^4 residue caused chemical shifts of the Gln-Val-Val-Ala-Gly sequence, suggesting that the N-terminal region might not be involved in direct interaction with papain, but may rather play a role in maintaining the conformation of the loop structure containing the Gln-Val-Val-Ala-Gly sequence (*24*). On the basis of these observations, it has been proposed that cystatins can be divided into two groups depending on the involvement of the N-terminal region in inhibitory activity (*24*): one group, including cystatin A, cystatin C, and chicken cystatin, has an N-terminal region whose conformation plays an important role in maintaining the first loop structure; the other group, including cystatin B, does not require the N-terminal region for inhibitory activity.

With regard to phytocystatins, Abe *et al.* constructed an oryzacystatin I mutant in which the N-terminal 21 amino acids were deleted and, instead, an extra 15 amino acids encoded by a pUC 18 polylinker were added, and showed that oryzacystatin I did not require the N-terminal region for papain inhibition (*25, 26*). However, it was reported that direct deletion of the N-terminal 21 amino acids in oryzacystatin caused a significant loss of papain-inhibitory activity (*27*). The present study showed that N-terminal truncation of Scb drastically decreased its inhibitory activity. In this regard, it is likely that Scb as well as oryzacystatin I can be classified into the former of the two

proposed groups; the N-terminal regions of both phytocystatins may support the conformation of the first loop as an intramolecular chaperone. By contrast, the other sunflower phytocystatin, Sca, which lacks 13 N-terminal amino acids as compared with Scb, would be classified into the latter group, as it can inhibit papain effectively (*8*). Although the functional implication of two types of cystatins is still not known, it is of interest that sunflower contains both types of phytocystatins in its seeds.

We thank Prof. Kenji Yamamoto and Dr. Tomoko Kadowaki, Faculty of Dentistry, Kyushu University, for assaying the inhibitory activity toward cathepsins B, L, and H.

REFERENCES

1. Abe, K. and Arai, S. (1985) Purification of a cysteine proteinase inhibitor from rice, *Oryza sativa* L. *japonica*. *Agric. Biol. Chem.* **49**, 3349-3350
2. Abe, K., Emori, Y., Kondo, H., Suzuki, K., and Arai, S. (1987) Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice seeds. *J. Biol. Chem.* **262**, 16793-16797
3. 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. Molecular cloning, expression, and biochemical studies on oryzacystatin-II. *J. Biol. Chem.* **265**, 15832-15837
4. 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
5. Abe, M., Abe, K., Kuroda, M., and Arai, S. (1992) Corn kernel cysteine proteinase inhibitor as a novel cystatin superfamily member of plant origin. Molecular cloning and expression studies. *Eur. J. Biochem.* **209**, 933-937
6. Misaka, T., Kuroda, M., Iwabuchi, K., Abe, K., and Arai, S. (1996) Soyacystatin, a novel cysteine proteinase inhibitor in soybean, is distinct in protein structure and gene organization from other cystatins of animal and plant origin. *Eur. J. Biochem.* **240**, 609-614
7. 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
8. Kouzuma, Y., Kawano, K., Kimura, M., Yamasaki, N., Kadowaki, T., and Yamamoto, K. (1996) Purification, characterization, and sequencing of two cysteine proteinase inhibitors, Sca and Scb, from sunflower (*Helianthus annuus*) seeds. *J. Biochem.* **119**, 1106-1113
9. Manning, K. (1991) Isolation of nucleic acids from plants by differential solvent precipitation. *Anal. Biochem.* **195**, 45-50
10. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
11. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
12. Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350-4354
13. Deng, W.P. and Nickoloff, J.A. (1992) Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal. Biochem.* **200**, 81-88
14. Von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* **133**, 17-21
15. Bode, W., Engh, 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
16. Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcie, B., and Turk, V. (1990) The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J.* **9**, 1939-1947
17. Thiele, U., Assfalg-Machleidt, I., Machleidt, W., and Auerswalt, E.A. (1990) N-terminal variants of recombinant stefin B: effect on affinity for papain and cathepsin B. *Biol. Chem. Hoppe-Seyler* **371**, 125-136
18. 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
19. Abrahamson, M., Mason, R.W., Hansson, H., Buttle, D., Grubb, A., and Ohlsson, K. (1991) Human cystatin C. Role of the N-terminal segment in the inhibition of human cysteine proteinases and in its inactivation by leucocyte elastase. *Biochem. J.* **273**, 621-626
20. Hall, A., Dalbøge, H., Grubb, A., and Abrahamson, M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem. J.* **291**, 123-129
21. Machleidt, W., Thiele, U., Laber, B., Assfalg-Machleidt, I., Esterl, A., Wiegand, G., Kos, J., Turk, V., and Bode, W. (1989) Mechanism of inhibition of papain by chicken egg white cystatin. Inhibition constants of N-terminally truncated forms and cyanogen bromide fragments of the inhibitor. *FEBS Lett.* **243**, 234-238
22. Shibuya, K., Kaji, H., Itoh, T., Ohyama, Y., Tsujikami, A., Tate, S., Takeda, A., Kumagai, I., Hirao, I., Miura, K., Inagaki, F., and Samejima, T. (1995) Human cystatin A is inactivated by engineered truncation. The NH_2-terminal region of the cysteine proteinase inhibitor is essential for expression of its inhibitory activity. *Biochemistry* **34**, 12185-12192
23. 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
24. Tate, S., Ushioda, T., Utsunomiya-Tate, N., Shibuya, K., Ohyama, Y., Nakano, Y., Kaji, H., Inagaki, F., Samejima, T., and Kainosho, M. (1995) Solution structure of a human cystatin A variant, cystatin $A^{2\text{-}98}$ M65L, by NMR spectroscopy. A possible role of the interactions between the N- and C-termini to maintain the inhibitory active form of cystatin A. *Biochemistry* **34**, 14637-14648
25. 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. Expression of oryzacystatin cDNA and its truncated fragments in *Escherichia coli*. *J. Biol. Chem.* **263**, 7655-7659
26. Arai, S., Watanabe, H., Kondo, H., Emori, Y., and Abe, K. (1991) Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J. Biochem.* **109**, 294-298
27. Urwin, P.E., Atkinson, H.J., and McPherson, M.J. (1995) Involvement of the NH_2-terminal region of oryzacystatin-I in cysteine proteinase inhibition. *Potein Eng.* **8**, 1303-1307
28. Hall, A., Håkansson, K., Mason, R.W., Grubb, A., and Abrahamson, M. (1995) Structural basis for the biological specificity of cystatin C. *J. Biol. Chem.* **270**, 5115-5121