Cloning, Functional Expression, and Characterization of Cystatin in Sesame Seed[†]

Douglas J. H. Shyu,[‡] Wing-Ming Chou,[§] Tien-Joung Yiu,[||]
Coney P. C. Lin,[‡] and Jason T. C. Tzen*,[‡]

Graduate Institute of Biotechnology, National Chung Hsing University,
Taichung, Taiwan 40227, Republic of China, Department of Biotechnology,
National Huwei Institute of Technology, Yunlin, Taiwan 63208, Republic of China, and
Crop Improvement Department, Tainan District Agricultural Research and Extension Station,
Tainan, Taiwan 70107, Republic of China

A cDNA fragment encoding cystatin, a cysteine protease inhibitor, was obtained from maturing sesame seeds. The clone was constructed in a nonfusion or fusion vector and then overexpressed in *Escherichia coli*. The recombinant cystatins were found in the soluble fraction of cell extract and were demonstrated to be functionally active in a reverse zymographic assay. The corresponding endogenous 22 kDa cystatin of low abundance in mature seeds was purified to homogeneity via a papain-coupling affinity column and confirmed by western blotting with antibodies against the recombinant cystatin. Both endogenous and recombinant cystatin proteins showed effective inhibitory activities against papain with K_i values of 7.89×10^{-8} M and 2.77×10^{-8} M, respectively. Immunodetection indicated that cystatin was specifically expressed in maturing seeds and rapidly degraded in germination. Accordingly, zymographic and inhibition analyses showed that sesame cystatin could not inhibit the de novo synthesized proteases in germinating seeds. It is suggested that sesame cystatin may play a role in the regulation of endogenous cysteine proteases during seed maturation and germination.

KEYWORDS: Cystatin; cysteine protease inhibitor; phytocystatin; seed; sesame

INTRODUCTION

Cystatins, a superfamily of cysteine protease inhibitors, have been investigated extensively in the past years (*1*, *2*). Cystatins of animal origin are classified into three families on the basis of their amino acid sequences, protein characteristics, and molecular structures (*2*, *3*). Family I, stefin, contains neither disulfide bond nor glycosylation site with molecular mass of approximately 11 kDa. Family II, cystatin, contains two disulfide bonds but no glycosylation site with molecular mass of approximately 15 kDa. Family III, kininogen, possesses three repeats of the cystatin proteins that are highly glycosylated. Kininogens usually contain extra sequences at their C-terminus with molecular masses around 60−120 kDa. In addition to these three well-established cystatin families, several proteins have been isolated that contain cystatin-like sequences without inhibitory activities (*4*).

In the past decade, several proteins homolgous to animal cystatins were isolated from plants and their corresponding

cDNAs were sequenced (*1*). It was found that those plant cystatins could not be categorized into the three known cystatin families. The size and sequence of plant cystatins are similar to those of family II cystatin but they contain neither disulfide bond nor cysteine residue. In all cystatins of animal and plant origins, three conservative regions are present including a glycine residue at the N-terminus, a QXVXG motif, and a tryptophan residue near the C-terminus (*2*, *3*). These three conserved regions have been demonstrated to directly interact with the active-site clefts of cysteine proteases in the papain family (*5*). In addition to the three conserved regions, plant cystatins possess a unique conserved sequence, LARF/YAV/IXXXN at the N-terminus of α1-helix, and thus are proposed to comprise a new subfamily termed phytocystatin (*3*).

Phytocystatins are assumed to regulate endogenous cysteine protease activities in response to biotic and abiotic stresses during seed development and germination, when stored reserves are mobilized for embryogenesis and seedling growth (*6*, *7*). They were also implicated in programmed cell death by modulating cysteine protease activity in the regulation of protein turnover (*8−10*). Phytocystatins may also protect plants from invasion of pathogens, nematodes, and insects (*1*). In addition, phytocystatins have been used in food processing and considered for other biotechnological applications (*1*, *11*).

Sesamum indicum Cystatin

J. Agric. Food Chem., Vol. 52, No. 5, 2004 **1351**

In this study, we obtained a cystatin cDNA clone from maturing sesame seeds while working on the molecular cloning of storage proteins (*12*). As cystatins play important roles in seed physiology and possess potential utilization in biotechnological applications, the recombinant proteins were overexpressed in *Escherichia coli* and the corresponding endogenous protein in mature sesame seeds was purified and characterized.

MATERIALS AND METHODS

Plant Materials. Mature and fresh maturing sesame (*Sesamum indicum* L.) seeds were grown in the Crop Improvement Department, Tainan District Agricultural Research and Extension Station. The seeds were soaked in water for 10 min prior to preparation of total seed proteins. Fresh maturing seeds 24 days after flowering were harvested for the construction of a cDNA library. For germination, mature seeds were imbided in water at 27 °C and harvested at different periods of time.

Protein Preparation from Various Tissues. Mature, maturing, and germinating seeds as well as other tissues including root, stem, leaf, flower, and pad were either used freshly or excised into small pieces, frozen in liquid nitrogen, and stored at −70 °C until used. The tissues were ground in a mortar with liquid nitrogen and extracted with a buffer containing 0.6 M sucrose and 0.01 M sodium phosphate, pH 7.5. The homogenate was filtrated through cheesecloth and subjected to protein analyses.

Cloning of Cystatin Gene. A cDNA library of approximately 10^6 plaques was constructed with mRNA extracted from maturing sesame seeds as described by Tai et al. (*12*). In an attempt to obtain other homologous genes, the clone encoding one of the 11S globulin precursor proteins (GenBank Accession Number AF091842) was isotope-labeled to screen the cDNA library, and the plaques showing relatively weak signals on hybridization were subjected to in vivo excision of the pBluescript phagemid from the Uni-ZAP XR vector. Potential clones were sequenced by the dideoxy chain-termination method with the BigDye terminator cycle sequencing kit and ABI 377 DNA sequencer (Perkin-Elmer). In addition to a homologous 11S clone, a full-length sequence (GenBank Accession Number AF240007) encoding a putative cystatin was identified via sequence comparison using the Blast program (*13*).

Overexpression of Recombinant Cystatin Proteins in *E. coli.* The sesame cystatin clone was constructed in the nonfusion expression vector pET28a(+) (Novagen) by use of *Nco*I sites. The recombinant vector was double-digested with *Nco*I and *Hin*dIII and constructed in the S-tag fusion expression vector pET29a(+) (Novagen). Both nonfusion and fusion constructs were separately transformed into *E. coli* BL21(DE3) and overexpressed under the control of T7 RNA polymerase/promoter system with the addition of 1 mM isopropyl thio-β-D-galactoside (IPTG). After induction for 3 h, the *E. coli* cells were harvested and lysed by sonication in 10 mM phosphate buffer, pH 8; the lysate was fractionated into soluble and precipitate fractions by centrifugation at 10000*g* and then subjected to SDS−PAGE and western blotting analyses.

Antibody Preparation, SDS−PAGE, and Western Blotting. Nonfusion recombinant cystatin overexpressed in *E. coli* was resolved on glycine−SDS−PAGE (*14*) and eluted according to the gel extraction method (*15*). Antibodies against the recombinant cystatin were raised in rabbits, and immunoglobulins were collected from serum for immunoassays (*16*). In immunoassays, proteins resolved by SDS−PAGE were transferred onto a nitrocellulose membrane in a Bio-Rad Trans-Blot system (Bio-Rad, Hercules, CA) according to the manufacturer's instruction. The membrane was subjected to immunodetection by use of secondary antibodies conjugated with horseradish peroxidase (Jackson ImmunoResearch Laboratories, Inc.) and then incubated with 4-chloro-1-naphthol (Sigma) containing hydrogen peroxide for color development.

Affinity Purification of the Recombinant and Endogenous Cystatin Proteins. Papain, a plant cysteine protease, has been used as a ligand enzyme to prepare an affinity column for cystatin purification (*17, 18*). Papain (Sigma) was coupled to CNBr-activated Sepharose 4B (Amersham Biosciences AB) at a concentration of 5 mg/mL of gel following the manufacturer's instruction. The nonfusion recombinant cystatin in the soluble fraction of *E. coli* cell lysate or endogenous cystatin in the soluble proteins extracted from mature sesame seeds was incubated with papain-coupled resins overnight at 4 °C. The cystatin-bound resins were equilibrated with 50 mM sodium phosphate buffer, pH 6.5, containing 0.5 M NaCl and 0.1% Brij 35, packed into a column, and washed with 50 mM sodium phosphate buffer, pH 6.5, containing 0.5 M NaCl and 10% (v/v) glycerol. Affiliated cystatin proteins were eluted with 50 mM K_3PO_4, pH 11.5, containing 0.5 M NaCl and 10% (v/v) glycerol. The eluent containing cystatin was adjusted to pH 7.4 with 5 M sodium formate buffer, pH 3.0, and concentrated by Amicon Ultra-15 (Millipore). Protein concentration was determined by the Bradford method (*19*).

Papain Inhibitory Activity Measurement. Cysteine protease inhibitory assay was performed according to the methods described by Barrett (*20*) and Abe et al. (*21*) with papain as the target enzyme and *N*-benzoyl-L-arginine-2-naphthylamide (BANA) (Sigma) as its substrate with some modifications. Papain (10 μg) was first incubated with various amounts of cystatin in 250 μL of assay solution containing 0.1 M sodium phosphate, pH 6, 1 mM EDTA, and 2 mM 2-mercaptoethanol at 37 °C for 5 min. Subsequently, proteolytic reaction was initiated by adding 0.1 mL of 1 mM BANA at 37 °C and terminated by adding 0.5 mL of 2% HCl/ethanol after reaction for 10 min. Color was developed by adding 0.5 mL of 0.06% *p*-(dimethylamino)-cinnamaldehyde (Sigma) in ethanol for 15 min, and the absorbance of the mixture was measured at 540 nm. Inhibitory activity was detected as a decreased amount of substrate-hydrolyzing activity of papain. The K_i values were determined by Dixon plots (*22*).

Reverse Zymography and Zymography. For reverse zymographic analysis, SDS−polyacrylamide gels were incubated in 10 mM Tris-HCl buffer, pH 7.5, containing 2.5% (v/v) Triton X-100 at room temperature for 30 min with gentle agitation. After two washings in 10 mM Tris-HCl buffer, pH 7.5, for 30 min, the gels were incubated with papain (10 μg/mL) in the protease activation buffer (0.1 M phosphate buffer, pH 6.0, containing 2 mM cysteine and 1 mM EDTA) at 37 °C for 1 h. The proteolytic reaction was stopped by transferring the gels to the SDS−PAGE staining solution. After gel staining and destaining, the visible blue bands represented active cysteine protease inhibitors. For zymographic analysis, proteins in an SDS−polyacrylamide gel were electrotransferred into another gel containing 0.1% (w/v) porcine gelatin. Electrotransfer was carried out at 4 °C in 25 mM Tris-HCl buffer, pH 8.3, containing 0.2 M glycine, under a constant voltage of 60 V for 20 min (*23*) in a similar manner as described for western blotting. After transfer, the gel was washed and the proteolytic reaction was performed as described for the reverse zymographic analysis without adding papain. For inhibitory assays, protease inhibitors were added to the protease activation buffer and allowed to react with target proteases for 10 min at room temperature prior to the proteolytic reaction.

RESULTS

Isolation and Characterization of a Sesame Cystatin cDNA Clone. A cystatin gene was accidentally cloned from a cDNA library of maturing sesame seeds. This cDNA fragment (GenBank Accession Number AF240007) comprises 904 nucleotides, consisting of a 15-nucleotide 5′-untranslated region, an open reading frame of 600 nucleotides, and a 289-nucleotide 3′-untranslated region. The deduced sequence comprises 199 amino acid residues encoding a 22 kDa polypeptide that contains no cysteine residue. Sequence alignment of sesame cystatin with other phytocystatins and a chicken egg white cystatin is shown in **Figure 1**. A conserved glycine residue, proposed to interact with the catalytic site of the papain family, and a conserved tryptophan residue, involved in contributing the three-point interaction between inhibitor and protease, were localized separately on both sides near the cystatin motif, QXVXG (*1−3*). In addition, a unique consensus sequence, [LVI]-[AGT]-

```
                                                                                ★
SiCYS  ------------------------------------------------------------MATLGGVH--DSN-S---  12
GmCYS  MRAL-TSSSSTFIPKRY-SFFFFLSIL-----FALRSSSGGCSEYHHHHAP-----------MATIGGLR--DSQ-GSQN  59
LeCYS  ----------MRVIRSRAILIVLF-------LVSAFGLSEQGKSGGFCSEE----------MATLGGVH--DSHGSSQN  50
IbCYS  ---------MRVRFFFYFAYFLITLVFFPSVTLQSYSGGHRQEATGFCGEEGEREDNLIRMATTTLLGGIS--DSA-SAEN  68
MdCYS  -----QMKSESYISLLIASFFVLLLPAF---SSAAEESSTPIGCGHTSHRN----------MATLGGVH--ESH-GAQN  58
OsC-3  MRVAATTRPASSSAAAPLPLFLLLAVAA--------AAAALFLVGSASLAMA----------GHVLGGAH--DAP-SAAN  59
AtCYS  --------------MMRSRFLLFIVFF---SLSLFISSLIASDLGFCNEE----------MALVGGVG--DVP-ANQN  48
RcCYS  -----------------------------------------------M----------ATVQGGVH--DSPQGTAN  17
BrCYS  -----------------------------------------------------MAMLGGVR--DVP-SNEN  15
OsC-1  -------------------------------------------MSSDG-----------GPVLGGV---EPV-GNEN  19
OsC-2  ------------------------------------------MAEEAQS----------HAREGGRHPRQPA-GREN  24
GgCYS  -----------MAGARGCVVLL-------------AAALMLVGAVLGSED----------RSRLLGA----PVP-VDEN  40


                                        ★★★
SiCYS  NPDTHS LARFAVDQHNT KENGLL--ELVRVVEAREQVVAGTLHHLVLEVLD-AGKKK----LYEAK-------------  71
GmCYS  SVQTEA LARFAVDEHNK KQNSLL--EFSRVVRTQEQVVAGTLHHLTLEAIE-AGEKK----LYEAK-------------  118
LeCYS  SDEIHS LAKFAVDEHNK KENAMI--ELARVVKAQEQTVAGKLHHLTLEVMD-AGKKK----LYEAK-------------  109
IbCYS  SVEIES LARFAVEEHNK KENAMI--ELVRVVKAEEQVVAGKLHHLTLEVID-AGKRK----LYEAK-------------  127
MdCYS  SAEVED LARFAVQEHNN KENALL--EFVSVVKAKEQVVAGTLHHLTIEAIE-AGKKK----LYQAK-------------  117
OsC-3  SVETDA LARFAVDEHNK RENALL--EFVRVVEAKEQVVAGTLHHLTLEAIE-AGKKK----VYEAK-------------  118
AtCYS  SGEVES LARFAVDEHNK KENALL--EFARVVKAKEQVVAGTLHHLTLEILE-AGQKK----LYEAK-------------  107
RcCYS  NAEIDG IARFAVDEHNK KENAMV--EFGRVLKAKEQVVAGTLHHLTIEAIE-AGKKK----IYEAK-------------  76
BrCYS  SVEVES LARFAVDEHNK KENALL--EFARVVKAKEQVVAGTMHHLTLEIIE-AGKKK----LYEAK-------------  74
OsC-1  DLHLVD LARFAVTEHNK KANSLL--EFEKLVSVKQQVVAGTLYYFTIEVKE-GDAKK----LYEAK-------------  78
OsC-2  DLTTVE LARFAVAEHNS KANAML--ELERVVKVRQQVVGGFMHYLTVEVKEPGGANK----LYEAK-------------  88
GgCYS  DEGLQRALQFAMAEYNRASNDKYSSRVVRVISAKRQLVSGIKYILQVEIGR-TTCPKSSGDLQSCEFHDEPEMAKYTTCT  119


            ★
SiCYS  --IWVKPWMDFKQLQEFKHVRD--V-------PSFTSSDLGAKTDDQVSGWRPVPVHDPVVQDAAHHAIKTIQERSNSLF  130
GmCYS  --VWVKPWLNFKELQEFKPAGD--V-------PSFTSADLGVKKDGHQPGWQSVPTHDPQVQDAANHAIKTIQQRSNSLV  187
LeCYS  --VWVKPWLNFKELQEFKHVED--V-------PTFTSSDLGVKQVEQNSGLKSVPVHDPVVEEAAEHAIKTIQQRSNSIH  178
IbCYS  --VWLKPWMNFKELQGFNHIED--I-------PTLTSSDLGAKRDWPNTGLKSVPVNDPVVQDEAAQHAVKTIQQRSNSLL  196
MdCYS  --VWVKPWMGFKEVQEFKHADEEET------PSVTSSDLGVKQGGHPPGWQSVPPHDPQVQDAANHAVKSLQQRSNSLL  188
OsC-3  --VWVKPWLDFKELQEFRNTGD--A-------TTFTNADLGAKKGGHEPGWRDVPVHDPVVKDAADHAVKSIQQRSNSLF  187
AtCYS  --VWVKPWLNFKELQEFKPASD--A-------PAITSSDLGCKQGEHESGWRVPGDDPEVKHVAEQAVKTIQQRSNSLF  176
RcCYS  --VWVKPWLNFKELQEFKHATD--VADTTASHPSFTSSDLGVKREGHGAEWKEVAAHDPVQDAATHAVNTIQQRSNSLF  152
BrCYS  --VWVKPWLNFKELQEFKPS-----------TTITPSDLGCKKGEGASGWREVPGDDPEVQHVADHAVKSIQQRSNSLF  140
OsC-1  --VWEKPWMDFKELQEFKPVDA--S-------ANA----------------------------------------  102
OsC-2  --VWERAWENFKQLQDFKPLDD--A-------TA-----------------------------------------  107
GgCYS  FVVYSIPWLNQIKLLESKCQ---------------------------------------------------------  139


SiCYS  PYELSEVVHANAEVVDTSAKFDMLLKVKRGGKEEKYKVEVHKSTEEGGFNLKKVDLDHS-----  199
GmCYS  PYELHEVADAKAEVIDDFAKFNLLLKVKRGQKEEKFKVEVHKNNQ-GGFHLNQMEQDHS-----  245
LeCYS  PYKLQEIVHANAEMADDSTKLHLVIKTSRGGKEEKFKVQVQHKNE-GAFHLNRMEPDN------  235
IbCYS  PYELQEIVHANAEVIDDSAKVHMLIKTKRGEKEEKFSVEVPKNNE-GAFFLNHMAPANS-----  254
MdCYS  PYELQEVVHAQAEVAEEHAKFNMLLKVKRGSKEEKFKAEVHKNME-GTFSLNQMEADHS-----  246
OsC-3  PYELLEIVRAKAEVVEDFAKFDILMKLKRGNKEEKFKAEVHKNLE-GAFVLNQMQQEHDESSSQ  250
AtCYS  PYELLEVVHAKAEVTGEAAKYNMLLKLKRGEKEEKFKVEVHKNHE-GALHLNHAEQHHD-----  234
RcCYS  PYQLQEIVHAKAQVVDDFAKFDMILKVKRGTSEEKFKVEVHKNNE-GTFLLNQMEPHT------  209
BrCYS  PYELQEVVHANAEVTGEAAKYNMVLKLKRGEKEEKFKVEVHKNHE-GVLHLNHMEQQHHD----  199
OsC-1  -----------------------------------------------------------  102
OsC-2  -----------------------------------------------------------  107
GgCYS  -----------------------------------------------------------  139
```

Figure 1. Sequence alignment of cystatins from sesame (SiCYS, AF240007), soybean (GmCYS, T07139), tomato (LeCYS, AF198388), sweet potato (IbCYS, AF241536), apple (MdCYS, AY173139), rice (OsC-1, -2, and -3, P09229, A38375, and AP001073), thale cress (AtCYS, AC069474), castor bean (RcCYS, T10057), mustard (BrCYS, S65071), and chicken egg white (GgCYS, P01038). The amino acid number for the last residue in each line is listed on the right for each protein. Broken lines in the sequences represent gaps introduced for maximum sequence similarity. Five conserved amino acid residues, including glycine (G) near the N terminus, glutamine (Q), valine (V) and glycine (G) at the central motif, and tryptophan (W) near the C terminus, are marked by stars. The boxed region indicates the unique conserved amino acid sequence LARF/YAV/IXXXN in the phytocystatin subfamily.

[RKE]-[FY]-[AS]-[VI]-X-[EDQV]-[HYFQ]-N, was found at the N-terminal proximity of sesame cystatin and other phytocystatins.

Interestingly, the extension sequences found at C-termini of the aligned phytocystatins, except oryzacystatins I and II, distinguished phytocystatins into two subgroups. The biological function of this C-terminal extension, which has no effect on the cysteine protease inhibitory activity, remains unknown at present (*24, 25*). Phylogenetic analyses support the presence of these two phytocystatin subgroups (data not shown). Nevertheless, molecular modeling of sesame cystatin could be built by use of the oryzacystatin I NMR structure (*5*) as a template, on the basis of their high homology (data not shown).

**Functional Expression of Sesame Cystatin in *E. coli*.** The cDNA fragment of cystatin was constructed in a nonfusion or a fusion expression vector and then overexpressed in *E. coli* BL21(DE3). As predicted by theoretical calculation, the fusion recombinant cystatin was 3 kDa larger than the nonfusion recombinant cystatin of 22 kDa (**Figure 2A**). These two recombinant proteins were predominantly present in the soluble fraction of *E. coli* cell lysate. Western blotting showed that both recombinant cystatins could be recognized by antibodies raised against the nonfusion recombinant cystatin (**Figure 2B**). Reverse zymographic assay indicated that the recombinant cystatins were functionally active and resistant to papain digestion in gel (**Figure 2C**).

Purification of Endogenous Cystatin from Mature Sesame Seeds. To purify endogenous cystatin, soluble proteins extracted from mature sesame seeds were subjected to papain-coupled affinity chromatography. A putative endogenous cystatin, 22

Sesamum indicum Cystatin

J. Agric. Food Chem., Vol. 52, No. 5, 2004 **1353**
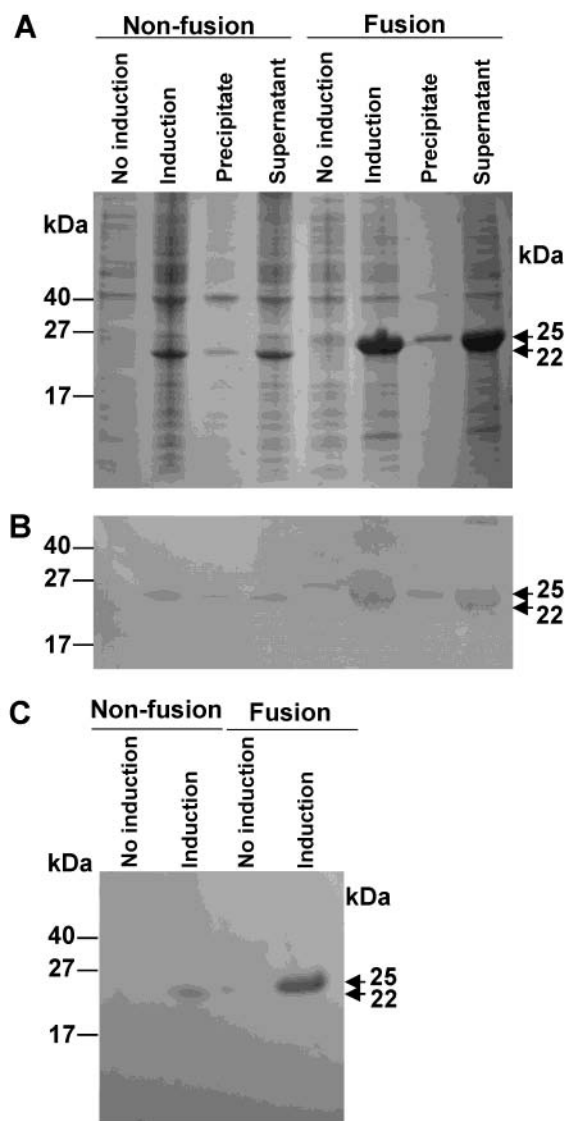
Figure 2. SDS−PAGE, western blotting, and reverse zymographic analysis of the overexpressed nonfusion and fusion recombinant cystatins in *E. coli.* (A) Total proteins were extracted from *E. coli* cells overexpressed in a nonfusion or fusion vector before or after IPTG induction, fractionated into precipitate and supernatant, and resolved by SDS−PAGE. (B) A duplicate gel was transferred onto a nitrocellulose membrane and then subjected to immunoblotting with antibodies against the nonfusion recombinant cystatin. (C) Recombinant cystatins in *E. coli* cells were subjected to reverse zymographic analysis. The visible bands represent papain inhibitors. Labels on the left indicate the molecular masses of three marker proteins, oleosin (17 kDa), caleosin (27 kDa), and stereoleosin (40 kDa), extracted from sesame oil bodies (*35, 36*). Two arrows on the right indicate the nonfusion and the fusion recombinant sesame cystatins of 22 and 25 kDa, respectively.

kDa estimated in SDS−PAGE, was purified to apparent homogeneity (**Figure 3**). Similarly, the nonfusion recombinant cystatin overexpressed in *E. coli* could also be purified to homogeneity via the same procedure. The putative endogenous cystatin and the nonfusion recombinant cystatin migrated correspondingly in SDS−PAGE and appeared to be the same protein according to western blotting. It is assumed that the cystatin clone obtained from the maturing seed cDNA library encodes the 22 kDa polypeptide affinity-purified from mature sesame seeds, and approximately 1 mg of endogenous cystatin is present in 1 kg of mature sesame seeds.
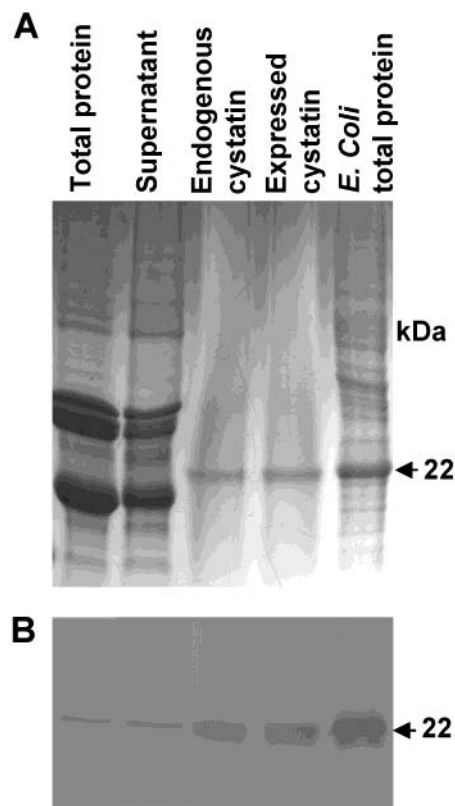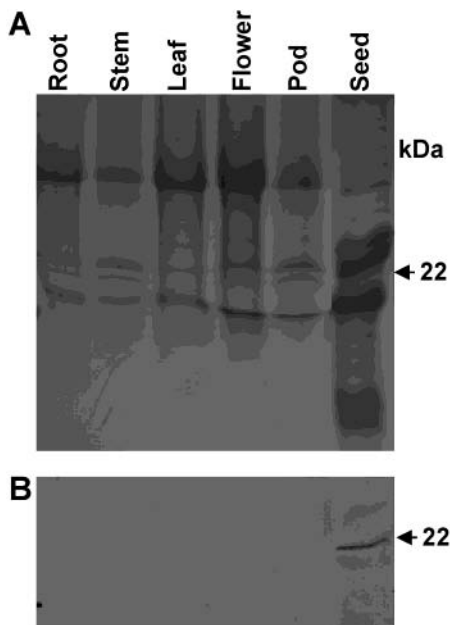

Figure 3. SDS−PAGE and western blotting of affinity-purified endogenous sesame cystatin and nonfusion recombinant cystatin. Endogenous cystatin was purified from soluble proteins extracted from mature sesame seeds, and nonfusion recombinant cystatin was purified from soluble proteins of *E. coli* lysate by a papain-coupled affinity column. The crude extracted proteins and purified cystatins were subjected to SDS−PAGE (A) and western blotting (B) analyses. Arrow on the right indicates the position of sesame cystatin (22 kDa).
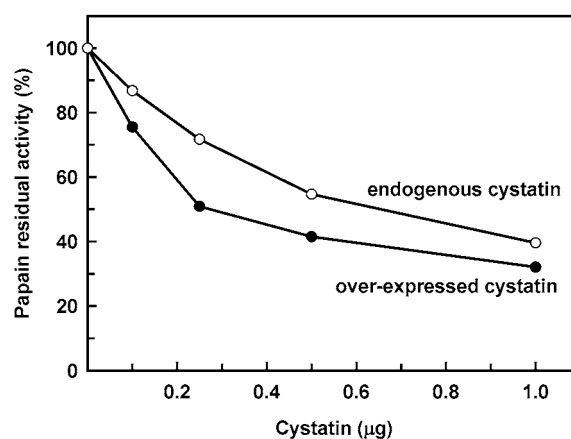

Figure 4. Inhibition of papain activity by endogenous and recombinant sesame cystatin. Papain was mixed with various amounts of the purified endogenous cystatin or nonfusion recombinant cystatin. The residual enzyme activities in the presence of cystatin were determined.

Inhibitory Activity of Sesame Cystatin against Papain. To compare the functional activity between endogenous and recombinant cystatins, both affinity-purified proteins were subjected to inhibitory activity analysis toward papain (**Figure 4**). Both endogenous and recombinant cystatins showed effective inhibitory activity against papain with K_i values of 7.89×10^{-8} M and 2.77×10^{-8} M, respectively. In comparison with other plant cystatins (*1*), which in the 10^{-8} M range strongly inhibit

Figure 5. SDS–PAGE and western blotting of cystatin in various sesame tissues. Total proteins were extracted from various tissues including root, stem, leaf, flower, pod, and seed. Samples of 50 μg were subjected to SDS–PAGE (A) and western blotting (B) analyses. Arrow on the right indicates the position of sesame cystatin (22 kDa).
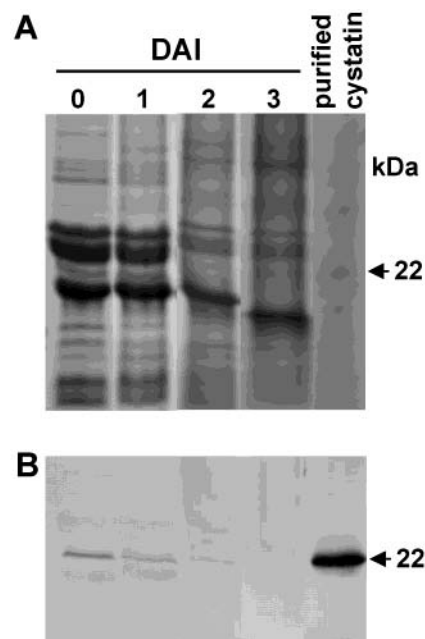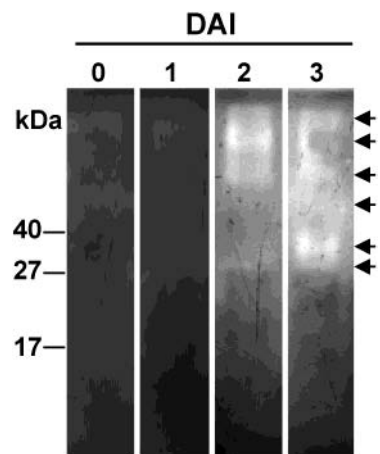

Figure 7. SDS–PAGE and western blotting of cystatin at different stages of germination. Total proteins extracted from germinating seeds at various days after imbibition (DAI) were subjected to SDS–PAGE (A) and western blotting (B) analyses. Arrow on the right indicates the position of sesame cystatin (22 kDa).
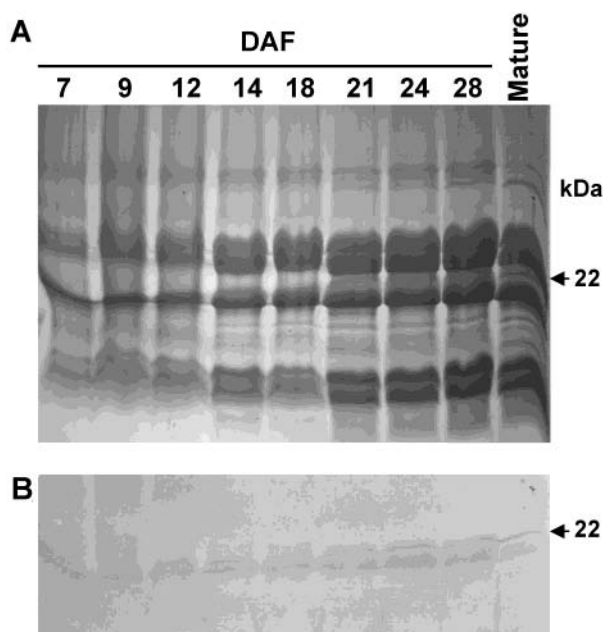
Figure 6. SDS–PAGE and western blotting of cystatin at different stages of maturing sesame seeds. Total proteins extracted from maturing sesame seeds at various days after flowering (DAF) were subjected to SDS–PAGE (A) and western blotting (B) analyses. Arrow on the right indicates the position of sesame cystatin (22 kDa).

Figure 8. Zymographic analysis of protease activity in different stages of germinating seeds. Total proteins extracted from germinating seeds at various days after imbibition (DAI) were resolved by SDS–PAGE, transferred to another gelatin-containing SDS–polyacrylamide gel, and then detected by the renaturation, proteolysis, and staining procedures as described under Materials and Methods. Potential proteases appeared as clear zones on the gel. Labels on the left indicate the molecular masses of three marker proteins, oleosin (17 kDa), caleosin (27 kDa), and stereoleosin (40 kDa), extracted from sesame oil bodies.

papain, the sesame cystatin with variable extension sequences at N- and C-terminus exhibits no significant decrease on its protease inhibitory activity, at least to papain.

Distribution of Cystatin in Sesame. According to western blotting analyses of proteins extracted from various tissues, cystatin was found to exclusively accumulate in seed (**Figure 5**). A detailed analysis showed that cystatin started to accumulate in maturing seeds approximately 14 days after flowering and

progressively accumulated till seed maturation (**Figure 6**); after germination it declined rapidly and disappeared 2 days after imbibition (**Figure 7**). Zymographic analysis revealed that several proteases of high molecular masses were de novo synthesized after seed germination (**Figure 8**). The occurrence of these proteases was concomitant with the decrease of endogenous cystatin shown in **Figure 7**. The results indicated that the de novo synthesized proteases may not be the inhibitory targets of endogenous cystatin. The inhibition experiment was thus performed to test this possibility. Accordingly, proteolytic activity of these postgerminatively synthesized proteases could

Sesamum indicum Cystatin

J. Agric. Food Chem., Vol. 52, No. 5, 2004 **1355**

not be inhibited in the presence of the endogenous or recombinant cystatin (data not shown).

DISCUSSION

In this study, a cDNA clone encoding a putative cystatin was obtained from maturing sesame seeds. Sequence alignment and phylogenetic analyses revealed that the deduced cystatin could be grouped in the phytocystatin subfamily of the cystatin superfamily. This sesame cystatin, either purified from mature seeds or expressed in *E. coli*, possessed inhibitory activity toward papain with K_i values similar to those of other phytocystatins studied previously (*1*). The extra sequences at the N- and C-termini adjacent to the core domain of the protein, though their biological function remains unresolved, seem to have no influence on the inhibitory activity, as reported previously for oryzacystatin and soyacystatin (*24, 25*). The extension sequences of cystatins were suggested to affect the affinities of these inhibitors for various endogenous or exogenous cysteine proteases.

Similar to oryzacystain (*24*), soyacystatin (*25*), and corn cystatin (*21*), sesame cystatin was found to be a seed-specific protein that started to accumulate in the middle stage of seed maturation. However, the destination of seed cystatin does not seem to gain as much attention as its accumulation. In this study, sesame cystatin was found to decline drastically after seed imbibition and could not be detected 2 days later. The zymographic studies revealed that several de novo synthesized proteases of high molecular masses could be detected after germination. It is speculated that the sesame cystatin was degraded by the de novo proteases during seed germination, thus relieving the inhibition of certain endogenous cysteine proteases, which are presumably involved in certain physiological functions of seedling growth, e.g., the conversion of storage proteins in protein bodies to free amino acids. In fact, diverse types of proteases identified within the plant vacuoles are responsible for the degradation and mobilization of proteins within protein bodies (*26*). Prevention of storage proteins from degradation during seed maturation may be assured by keeping the proteases in an inactive state and/or by modifying storage proteins in an inaccessible structure. It was reported that the proteases of dicotyledonous plants remained inactive before de novo synthesized proteases entered the protein bodies (*7, 27*). On the other hand, it was also reported in vetch that globulin mobilization was initiated by stored proteases during germination, and the de novo synthesized enzymes mediated the bulk breakdown of stored globulins after germination (*28, 29*).

Plant cystatins, in addition to their biological function as endogenous enzyme inhibitors, may impede exogenous target enzymes as well (*1*). Recent research has extensively focused on the roles of cystatins in controlling crop invaders, whose digestive systems are mainly composed of cysteine proteases. As insects utilize multiple proteases to facilitate nutrient assimilation, feeding insect pests with serine protease inhibitors only partly prevents crops from damage. Therefore, possession of a combination of two or more protease inhibitors may endow a plant with significant pest-resistance effects (*30*). Moreover, cystatins have also been implicated in plant responses to a variety of biotic and abiotic stresses (*31, 32*) and thus may play crucial roles in the defense mechanism. For the study of plant phenotype−environment relationship, effects of transgenic plants expressing cystatins are being investigated intensively (*33*).

Recently, the potential use of cystatins as food additives to prevent cysteine proteases from degrading structural proteins during meat and surimi processing has been investigated (*34*). Phytocystatins have gained attention in these potential applications due to their effective inhibitory activity and a lack of requirement for formation of disulfide bonds. Unfortunately, difficulty was found with low yields of cystatin proteins when directly purified from plant materials, leading to a high cost of production. In the present study, we have isolated a cystatin cDNA clone from sesame, and its overexpression in *E. coli* produced remarkably soluble and active proteins. More biotechnological applications of this remunerative recombinant cystatin will be evaluated in our follow-up studies.

ABBREVIATIONS USED

BANA, N-benzoyl-L-arginine-2-naphthylamide; CNBr, cyanogen bromide; IPTG, isopropyl thio-β-D-galactoside

LITERATURE CITED

(1) Arai, S.; Matsumoto, I.; Emori, Y.; Abe, K. Plant seed cystatins and their target enzymes of endogenous and exogenous origin. *J. Agric. Food Chem.* **2002**, *50*, 6612−6617.

(2) Turk, B.; Turk, V.; Turk, D. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol. Chem.* **1997**, *378*, 141−150.

(3) Margis, R.; Reis, E. M.; Villeret, V. Structural and phylogenetic relationships among plant and animal cystatins. *Arch. Biochem. Biophys.* **1998**, *359*, 24−30.

(4) Brown, W. M.; Dziegielewska, K. M. Friends and relations of the cystatin superfamily−new members and their evolution. *Protein Sci.* **1997**, *6*, 5−12.

(5) Nagata, K.; Kudo, N.; Abe, K.; Arai, S.; Tanokura, M. Three-dimensional structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. *japonica*. *Biochemistry* **2000**, *39*, 14753−14760.

(6) Abe, K.; Emori, Y.; Kondo, H.; Suzuki, K.; Arai, S. Molecular cloning of a cysteine proteinase inhitor of rice (oryzacystatin)−Homology with animal cystatins and transient expression in the ripening process of rice seed. *J. Biol. Chem.* **1987**, *262*, 16793−16797.

(7) Callis, J. Regulation of protein degradation. *Plant Cell* **1995**, *7*, 845−857.

(8) Solomon, M.; Belengh, B.; Delledonne, M.; Levine, A. The involvement of cysteine proteases and protease inhibitor genes in programmed cell death in plants. *Plant Cell* **1999**, *11*, 431−444.

(9) Xu, F.-X.; Chye, M.-L. Expression of cysteine proteinase during developmental events associated with programmed cell death in brinjal. *Plant J.* **1999**, *17*, 321−327.

(10) Belenghi, B.; Acconcia, F.; Trovato, M.; Perazzolli, M.; Bocedi, A.; Polticelli, F.; Ascenzi, P.; Delledonne, M. *AtCYS1*, a cystatin from *Arabidopsis thaliana*, suppresses hypersensitive cell death. *Eur. J. Biochem.* **2003**, *270*, 2593−2604.

(11) Gutierrez-Campos, R.; Torres-Acosta, J. A.; Saucedo-Arias, L. J.; Gomez-Lim, M. A. The use of cysteine proteinase inhibitors to engineer resistance against potyviruses in transgenic tobacco plants. *Nat. Biotechnol.* **1999**, *17*, 1223−1226.

(12) Tai, S. S. K.; Lee, T. T. T.; Tsai, C. C. Y.; Yiu, T.-J.; Tzen, J. T. C. Expression pattern and deposition of three storage proteins, 11S globulin, 2S albumin and 7S globulin in maturing sesame seeds. *Plant Physiol. Biochem.* **2001**, *39*, 981−992.

(13) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389−3402.

(14) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.

1356 *J. Agric. Food Chem.,* Vol. 52, No. 5, 2004

Shyu et al.

(15) Tzen, J. T. C.; Chuang, R. L. C.; Chen, J. C. F.; Wu, L. S. H. Coexistence of both oleosin isoforms on the surface of seed oil bodies and their individual stabilization to the organelle. *J. Biochem. (Tokyo)* **1998**, *123*, 319−324.

(16) Chen, E. C. F.; Tai, S. S. K.; Peng, C.-C.; Tzen, J. T. C. Identification of three unique proteins in seed oil bodies of sesame. *Plant Cell Physiol.* **1998**, *39*, 935−941.

(17) Anastasi, A.; Brown, M. A.; Kembhavi, A. A.; Nicklin, M. J. H.; Sayers, C. A.; Barrett, A. J. Cystatin, a protein inhibitor of cysteine proteinases. *Biochem. J.* **1983**, *211*, 129−138.

(18) Benjakul, S.; Visessanguan, W.; An, H. Properties of cysteine proteinase inhibitors from black gram and rice bean. *J. Food Biochem.* **2001**, *25*, 211−227.

(19) Bradford, M. M. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein−dye binding. *Anal. Biochem.* **1976**, *72*, 248−254.

(20) Barrett, A. J. A new assay for cathepsin B1 and other thio proteinases. *Anal. Biochem.* **1972**, *47*, 280−293.

(21) Abe, M.; Abe, K.; Iwabuchi, K.; Domoto, C.; Arai, S. Corn cystatin I expressed in *Escherichia coli*: investigation of its inhibitory profile and occurrence in corn kernels. *J. Biochem. (Tokyo)* **1994**, *116*, 488−492.

(22) Dixon, M. The determination of enzyme inhibitor constants. *Biochem. J.* **1953**, *55*, 170−171.

(23) Visal-Shah, S.; Vrain, T. C.; Yelle, S.; Nguyen-Quoc, B.; Michaud, D. An electroblotting, two-step procedure for the detection of proteinases and the study of proteinase/inhibitor complexes in gelatin-containing polyacrylamide gels. *Electrophoresis* **2001**, *22*, 2646−2652.

(24) Abe, K.; Emori, Y.; Kondo, H.; Arai, S.; Suzuki, K. The NH₂-terminal 21 amino acid residues are not essential for the papain-inhibitory activity of oryzacystatin, a member of the cystatin superfamily. *J. Biol. Chem.* **1988**, *262*, 7655−7659.

(25) Misaka, T.; Kuroda, M.; Iwabuchi, K.; Abe, K.; Arai, S. 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.* **1996**, *240*, 609−614.

(26) Vierstra, R. D. Proteolysis in plants: mechanisms and functions. *Plant Mol. Biol.* **1996**, *32*, 275−302.

(27) Müntz, K. Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seeds. *J. Exp. Bot.* **1996**, *47*, 605−622.

(28) Müntz, K.; Belozersky, M. A.; Dunaevsky, Y. E.; Schlereth, A.; Tiedemann, J. Stored proteinases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *J. Exp. Bot.* **2001**, *52*, 1741−1752.

(29) Schelereth, A.; Standhardt, D.; Mock, H.-P.; Müntz, K. Stored cysteine proteinases start globulin mobilization in protein bodies of embryonic axes and cotyledons during vetch (*Vicia sativa* L.) seed germination. *Planta* **2001**, *212*, 718−727.

(30) Jouanain, L.; Bonadé-Bottino, M.; Girard, C.; Morrot, G.; Giband, M. Transgenic plants for insect resistance. *Plant Sci.* **1998**, *131*, 1−11.

(31) Botella, M. A.; Xu, Y.; Prabha, T. N.; Zhao, Y.; Naraimhan, M. L.; Wilson, K. A.; Nielsen, S. S.; Bressan, R. A.; Hasegawa, P. M. Differential expression of soybean cysteine proteinase inhibitor genes during development and in response to wounding and methyl jasmonate. *Plant Physiol.* **1996**, *112*, 1201−1210.

(32) Pernas, M.; Sanchez-Mong, R.; Salcedo, G. Biotic and abiotic stress can induce cystatin expression in chestnut. *FEBS Lett.* **2000**, *467*, 206−210.

(33) Van der Vyver, C.; Schneidereit, J.; Driscoll, S.; Turner, J.; Kunert, K.; Foyer, C. H. Oryzacystatin I expression in transformed tobacco produces a conditional growth phenotype and enhances chilling tolerance. *Plant Biotechnol. J.* **2003**, *1*, 101−112.

(34) Jiang, S.-T.; Tzeng, S.-S.; Wu, W.-T.; Chen, G.-H. Enhanced expression of chicken cystatin as a thioredoxin fusion form in *Escherichia coli* AD494(DE3)pLysS and its effect on the prevention of surimi gel softening. *J. Agric. Food Chem.* **2002**, *50*, 3731−3737.

(35) Chen, J. C. F.; Tsai, C. C. Y.; Tzen, J. T. C. Cloning and secondary structure analysis of caleosin, a unique calcium-binding protein in oil bodies of plant seeds. *Plant Cell Physiol.* **1999**, *40*, 1079−1086.

(36) Lin, L.-J.; Tai, S. S. K.; Peng, C.-C.; Tzen, J. T. C. Stereoleosin, a sterol-binding dehydrogenase in seed oil bodies. *Plant Physiol.* **2002**, *128*, 1200−1211.