J. Biochem. 126, 769-775 (1999)

Molecular Cloning, Enhancement of Expression Efficiency and Site-Directed Mutagenesis of Rat Epidermal Cystatin A[1]

Hiroyuki Kaji,* Rumi Yada-Wakatabe,† Takashi Uehira,† Mayumi Terai,† Atsushi Takeda,‡ Takanori Satoh,† and Tatsuya Samejima†,[2]

*Department of Chemistry, Graduate School of Science, Tokyo Metropolitan University, Minami-osawa, Hachioji-shi, Tokyo 192-0397; †Department of Chemistry, College of Science and Engineering, Aoyama Gakuin University, Chitosedai, Setagaya-ku, Tokyo 157-8572; and ‡Faculty of Arts and Sciences, Sagami Women's University, Bunkyo, Sagamihara-shi, Kanagawa 228-8533

Received May 7, 1999; accepted August 6, 1999

A rat cystatin A cDNA clone was isolated from a λ ZAP library representing newborn rat skin mRNA by screening with a synthetic oligonucleotide designed from amino acid sequence 15-23 of the cysteine proteinase inhibitor. The obtained clone contained a partial coding region of the inhibitor, lacking the 5′-untranslated region and coding sequence for the NH_2-terminal 13 residues. The amino acid sequence deduced from the base sequence, Glu14-Phe103, coincided with that determined at the amino acid level. To obtain the recombinant cystatin A protein, the DNA was fused with a synthetic linker encoding its missing N-terminal 17 residues and introduced into an expression vector, pMK2. In *Escherichia coli*, however, the expression level of the semi-synthetic gene was low, 0.5 mg of the purified recombinant protein per 1 liter culture being produced. Changing of the codon usage of the N-terminal region in a pET-15b expression system led to an increase in the yield depending on the instability of the putative secondary structure around an initiation codon of the mRNA. The expressed cystatin A showed identical characteristics with the authentic form except for the absence of the N-terminal acetyl blocking group. Using the expression system, two kinds of point mutation, the conservative Val54 in the first loop QxVxG region being changed to Lys and Glu, were introduced, but there was almost no effect on the inhibitory activity toward papain. This suggests that the conserved Val in the reactive site is not restricted and that the hydrophobicity of the position is not essential for the activity of rat cystatin A.

Key words: cystatin, expression efficiency, rat, secondary structure of mRNA, site-directed mutagenesis.

Rat cystatin A, also called cystatin α, is an acidic proteinous cysteine proteinase inhibitor belonging to the cystatin superfamily (*1*). The members of this superfamily can bind tightly with a series of papain-like cysteine proteinases, such as mammalian lysosomal cathepsins B, H, L, and S, thereby inhibiting their enzymatic activity in a reversible manner. Cystatin A, a member of family 1 (stefins) of this superfamily, is located intracellularly and contains no disulfide bond or sugar, in contrast to extracellular family 2 members (cystatins) with disulfide bridges, and family 3 plasma proteins comprising kininogens. These inhibitors seem to be related with protection against unwanted proteolysis; however, a regulatory function of cystatin B has been suggested, *i.e.*, a defect of the cystatin B gene causes progressive myoclonus epilepsy (*2*).

Rat cystatin A was isolated first from epidermis (*3, 4*), and was found to be localized specifically in epidermis, tongue, esophagus, and stomach, in contrast to ubiquitously distributed cystatin B (β) of the same family (*5*). In epidermis, it was proposed that cystatin A becomes located in keratohyalin granules as a phosphorylated form and then becomes cross-linked with suitable protein(s) through transglutaminase in the presence of Ca^{2+} to form a cornified cell envelope (*6-8*). Recently, human epidermal cystatin A was reported to be identical with keratolinin, which is one of the precursor proteins of the cornified cell envelope (*9*). In cultured keratinocytes, however, no cystatin A fragment was detected in the cornified envelope (*10*). The expression of the gene was induced by Ca^{2+} and TPA, a classical protein kinase C activator (*11*), and by UV irradiation (*12*). On the other hand, mouse cystatin A was cloned as a skin Ca^{2+}-binding protein, and its expression was regulated on differentiation and tumorigenesis (*13*). Similarly, chicken cystatin was described as a phosphorylated protein in part (*14*), and could bind Ca^{2+} as well as terbium (*15*). These results suggest that the significance of cystatin A is not only in its protective role but also in its mechanical support of the epidermis. However, the physiological function related

[1] This work was supported partly by a Grant-in-Aid for Scientific Research on Priority Areas (Intracellular Proteolysis) from the Ministry of Education, Science, Sports and Culture of Japan.

[2] To whom correspondence should be addressed. Tel: +81-3-5384-1111 (Ext. 23204), Fax: +81-3-5384-6200, E-mail: samejima@candy.chem.aoyama.ac.jp

Abbreviations: TPA, 12-*O*-tetradecanoylphorbol-13-acetate; PCR, polymerase chain reaction; IPTG, isopropyl-β-D-thiogalactopyranoside; CD, circular dichroism; API, *Achromobacter* protease I.

to the proteinase inhibitory action of cystatins remains unclear.

Sequence determination of newborn rat epidermal cystatin A revealed that it comprises 103 amino acid residues with an *N*-acetyl blocking group (*16*), and is accompanied by the presence of N-terminally truncated forms lacking the first 6 and 15 residues, respectively (*17*). Comparison of their inhibitory activities toward papain showed that the 6-residue truncated derivative was less active than the full-length inhibitor, and the 15-residue truncated form was completely inactivated. Therefore, the N-terminal region of cystatin A is essential for its inhibitory action. In fact, the N-terminal region and two conservative loops construct a tripartite wedge-shaped reactive site of cystatins, as found in a series of detailed three-dimensional structure analyses of chicken cystatin (*18-20*), human cystatin B complexed with papain (*21*), human cystatin A (*22, 23*), and human cystatin C (*24*). Judging from the results of many analyses of deletion and substitution mutants of cystatins, the main contribution to proteinase inhibition seems to be provided by the conservative QxVxG region in the first loop, and the N-terminal region around a conserved Gly might be responsible for the specificity toward target proteinases, regulation of their inhibitory activity, and structural integrity of the proteins. But the necessity of the N-terminal region is dependent on the molecular species, that is, a few cystatins need no conserved Gly residue for inhibition, as exemplified by human cystatin B (*25*) and rice seed oryzacystatin (*26*). Thus, it is still a problem as to what is an essential element in the reactive site of cystatins required for the tight binding with proteinases. To resolve this a problem and also the structure-function relations of cystatins, cDNA cloning, construction of expression plasmids, improvement of the expression level in *Escherichia coli* by changing the codon usage, and site-directed mutagenesis of rat cystatin A are described in this paper.

MATERIALS AND METHODS

Materials and Chemicals—Newborn rats, 2-day-old female Sprague Dawley, purchased from Nihon Clea Lab., were used for the preparation of cystatin A and mRNA. The messenger RNA fraction was purified on oligo(dT)-cellulose from Sigma, and a cDNA library was constructed using a λ ZAP II cDNA synthesis kit and an *in vitro* packaging kit, Giga Pack Gold, from Stratagene. Oligodeoxyribonucleotides were synthesized with the DNA synthesizers from Applied Biosystems, (381A) or MilliGen/Biosearch (Cyclone Plus). Restriction endonucleases, T4-polynucleotide kinase, T4-DNA ligase, and other DNA modifying enzymes were purchased from Nippon Gene, Toyobo, or Boehringer. [γ-^{32}P]ATP (220 TBq/mmol) and [^{35}S]dCTP(αS) (22 TBq/mmol) were from Amersham. Papain Type IV and the substrate, N^{a}-benzoyl-DL-arginine-2-naphthylamide (BANA), were from Sigma. The expression vector, pET-15b, was obtained from Novagen. DEAE-Sephacel and CNBr-activated Sepharose for affinity columns were from Pharmacia. Antiserum against rat cystatin A was prepared by the method of Vaitukaitis *et al.* (*27*). All other chemicals were of the highest grade available.

Preparation of Rat Epidermal Cystatin A—Rat cystatin A was prepared from newborn rat epidermis by the method of Takeda *et al.* (*28*) as reported previously. Briefly, epidermal cells were homogenized in 0.1 M Tris-HCl (pH 8.0) containing 0.145 M NaCl and 50 mg/liter phenylmethylsulfonyl fluoride, and then the resulting extract was freeze-dried after dialysis against distilled water. The lyophilized powder was dissolved in the buffer and then the supernatant was applied on a Sephadex G-75 column, followed by DEAE-Sephacel column chromatography. Cystatin A was monitored as to its inhibitory activity toward papain using BANA as a substrate according to the method described previously (*29*). The homogeneity of the isolated inhibitor was confirmed by SDS-polyacrylamide gel electrophoresis.

Construction and Screening of a Rat Skin cDNA Library—Rat skin was stripped from 2-day-old rats and frozen immediately in liquid nitrogen. The total RNA fraction was prepared by the PhOH/SDS method (*30*) from about 30 g of the skin, which was powdered in a stainless mortar in liquid nitrogen. Poly(A)$^{+}$-RNA was isolated by oligo(dT)-cellulose column chromatography. Double-stranded cDNA was synthesized basically by the method of Gubler and Hoffman (*31*) according to the procedure recommended by the supplier, Stratagene. The resulting cDNA was ligated to the bacteriophage λ ZAP II vector predigested with *Eco*RI and *Xho*I, and then packaged *in vitro*. The cDNA library was amplified once in *E. coli* PLK-F′ and then stored in SM buffer (50 mM Tris-HCl (pH 7.5) containing 0.1 M NaCl, 0.2% $MgCl_2$, and 0.1% gelatin) with 5% chloroform at 4°C.

The cDNA library was screened using a ^{32}P-labeled synthetic oligonucleotide probe of 27 bases corresponding to amino acid residues Ala15-Gln23 of rat cystatin A, 5′-CTGGATTTCTGGTGTGGCAGGTTTGGC-3′ of the antisense sequence. The sequence was designed according to the nucleotide sequences of cDNA of cystatin A from human epidermal keratinocytes (*12*) and mouse epidermis (*13*), because the amino acid sequence of the chosen area of rat cystatin A was identical with those of man and mouse, and the base sequences were quite homologous.

Recombinant plaques were formed on NZY-plates (0.5% yeast extract, 1% NZ-amine, 0.2% $MgCl_2$, and 0.5% NaCl) at a density of 50,000 plaques/100 cm^2, and then transferred to nylon membrane filters, Hybond N, Amersham. After the phage DNA had been fixed to the membranes, the filters were pre-hybridized in a solution containing 6×SSC, 5×Denhaldt's solution and 0.1% SDS at 37°C for 1 h. The probe oligomer was labeled with [γ-^{32}P]ATP with T4-polynucleotide kinase. The filters were hybridized in the same fresh solution with the labeled probe at 45°C overnight. After washing at room temperature, the filters were exposed to Fuji RX films using an intensifying screen at −80°C for 2 days. The positive plaques were selected and further screened under the same conditions. Positive plaques revealed on the secondary screening were analyzed again by Southern hybridization using the same probe.

Double-stranded cDNA inserts from the positive phages were subcloned into a phagemid vector, pBluescript, *in vivo*. *E. coli* XL1-Blue was infected with each positive λ phage and filamentous R408 helper phage simultaneously to excise the cDNA insert from the phagemid. *E. coli* was transfected again with the resulted phage harboring the phagemid and then formed transformant colonies exhibiting ampicillin-resistance. Double-stranded phagemid DNA

from the *E. coli* cells and single-stranded DNA from the amplified particles were sequenced using Sequenase version 2 with [^{35}S]dCTP(αS) and a Hitachi Fluorescence DNA Sequencer DSQ-3000, respectively.

Construction of an Expression Vector for Rat Cystatin A—Because the obtained cDNA insert lacked the partial coding region corresponding to upstream from Ser13 of cystatin A, a chemical synthetic linker for the missing region was introduced to construct an expression plasmid for the whole molecule of the inhibitor. Insertion cDNA was excised by *Eco*RI (upstream edge) and *Xho*I (down stream edge) digestion, and then subcloned once into the *Eco*RI-*Sal*I site of pUC18. To fuse the linker with the cDNA, a restriction site, *Nae*I, was introduced into the cDNA insert by site-directed mutagenesis according to the method of Morinaga *et al.* (*32*). A *Nae*I site was created at the position of Lys16-Ala18 (AAACCCGCC→AAgCCgGCC, substituted bases are indicated in lower case letters). The nucleotide sequence of the linker synthesized is shown in Fig. 1A. The insert was cut out by *Nae*I (created)-*Pst*I (multiple cloning site, outside of *Sal*I) digestion, and the fragment was ligated into the *Eco*RI (downstream of Shine-Dalgarno sequence)-*Pst*I site of a vector, pMK2 (*33*), with the linker. The resulting expression plasmid for rat cystatin A was named pRC81. The expression of the semi-synthetic gene in *E. coli* was confirmed by SDS-polyacrylamide gel electrophoresis of the cellular proteins. The transformant *E. coli* cells were cultured in Luria-Bertani medium containing 0.1 mg/ml of ampicillin (LB/Amp) at 37°C for 20 h. After denaturing gel electrophoresis on the *E. coli* cell lysate, the separated proteins were transferred to a nitrocellulose membrane (Hybond C; Amersham) and then stained immunologically with rabbit anti-rat cystatin A antiserum and alkaline phosphatase-conjugated goat anti-rabbit IgG.

Construction of High Efficiency Expression Vectors—To increase the expression level of rat cystatin A in *E. coli*, the expression vector was changed from pMK2 to pET-15b. The coding region of pRC81 was excised by digestion with *Pst*I followed by Klenow treatment and *Kpn*I digestion, as a *Kpn*I-Blunt ended fragment. An expression vector, pET-15b, was also digested by sequential treatments with *Bam*HI, Klenow and *Nco*I. The coding fragment was introduced into there with a synthetic linker corresponding to initial Met(*Nco*I)-Thr5(*Kpn*I), which comprised 5′-CATG-GACCCGGGTAC-3′ and 5′-CCGGGTC-3′. The resulting expression plasmid was named pET-CD. Furthermore, the codon usage for the amino-terminal region was changed randomly by PCR using a pair of primers having a mixed sequence: the sense-primer was 5′-AGAATTCCATGGAC-CCGGGNACNACNGGNATCGTNGGTGGTG-3′ (N indicates a mixture of four bases) and the antisense-primer was 5′-TTCTCGAGTCCGCCAAAACAGAAGCTTG-3′, corresponding to downstream of the insert in pMK2. The amplified fragment was digested with *Nco*I-*Xho*I and then inserted between the same sites of pET-15b. *E. coli* BL21(DE3) was transformed with the ligated DNA and then several clones showing relatively high expression efficiency for the gene were selected through SDS-PAGE of their cell lysates. These plasmids were designated as pET-Cr with the clone number. Each insert of the selected clones was subcloned into pUC118 as a *Xba*I-*Hin*dIII fragment to determine its sequence.

The minimum free energy for the secondary structure of each mRNA was calculated using the computer software, GENETYX, where transcription was supposed to be start at the second A of GGGGAATTGT in the *lac* operator sequence and the calculation was carried out for various positions of the transcripts; 1-100, 1-110, 1-120, 51-100, 51-110, and 51-120. The expression level of each clone was estimated by image analysis of a CBB-stained gel of total cellular proteins using the computer software, NIH-image.

Site-Directed Mutagenesis—A single amino acid substitution of Val 54 (GTC) to Glu (GAA) or Lys (AAA) was introduced by PCR with mutated primers. Two overlapping fragments were amplified by PCR separately using pET-Cr:3-1 as the template; an upstream fragment from the *T7* promoter to Leu 59 was amplified with 5′-CGATCCCGCG-AAATTAATACGACTCACT (sense primer) and 5′-AAgA-TcTGTCCAGCttnGACTTGAGATT (antisense, lower case letters indicate mutated points; AgATcT was used to create a new *Bgl*II site and ttn could change the codon to AAA (Lys), GAA (Glu), CAA (Gln), or TAA(termination)). Likewise, a downstream fragment from Gln 52 to the cloning site of the pET vector was amplified with 5′-CAAG-TCnaaGCTGGACAgATcTT (sense) and 5′-ATCCTCGAG-TCCGCCAAAACAGAAGCTTC (antisense). The resulting two fragments were mixed and subjected to further PCR with an upstream sense-primer and a downstream antisense-primer to form a full-length fragment. The secondary PCR product was digested with *Xba*I-*Pst*I and then subcloned into pUC118. After confirmation of its nucleotide sequence, this fragment was replaced with the same fragment of the original plasmid, pET-Cr:3-1, to construct mutant expression plasmids, pET-Cr:3-1(V54E) and (V54K) for Glu and Lys, respectively. The fragment replacement of each plasmid was confirmed by digestion with co-mutated *Bgl*II.

Purification and Characterization of Recombinant Rat Cystatin A—*E. coli* JM109 transformed with pRC81 was cultured in LB/Amp at 37°C for 20 h. *E. coli* cells were harvested by centrifugation, and then washed with 10 mM Tris-HCl (pH 8.0) containing 2 mM EDTA, 0.1 mg/ml phenylmethylsulfonyl fluoride, and 5% (v/v) EtOH. The cells were lyzed by sonication in the same buffer and then the supernatant was dialyzed against 10 mM Tris-HCl (pH 8.0) containing 5% EtOH. The soluble fraction was applied to a DEAE-Sephacel column and eluted with a linear gradient of NaCl, from 0 to 100 mM, in the same buffer. Cystatin A was detected as to the inhibitory activity as described earlier. The active fraction was subjected to *S*-carboxymethylated-papain Sepharose affinity chromatography followed by FPLC on a mono-Q column. In the advanced expression system, *E. coli* BL21(DE3) was used as host cells. A series of transformants was cultured in LB/Amp at 37°C and the expression was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at the mid-log phase of growth. After overnight cultivation, the cells were collected, and the cystatin A expressed was purified by DE-52 (Whatman) anion exchange chromatography followed by Sephadex G-75 (Pharmacia) gel filtration and then HPLC on a Shodex DEAE-825 column (Showa Denko). The homogeneity of the inhibitor preparations was confirmed by SDS-polyacrylamide gel electrophoresis. The amino-terminal amino acid sequence of the isolated intact inhibitor was determined with an automated sequencer from

Shimadzu PSQ-1. Peptide maps of both natural and recombinant cystatins were compared on reverse phase HPLC (RPC-C18S, JASCO) after digestion with *Achromobacter* protease I (Lysyl endopeptidase; Wako Pure Chemicals) under denaturation conditions with 3 M urea. A far UV circular dichroism (CD) spectrum was measured at room temperature with a JASCO recording dichrograph J-600. The inhibitory activity of the cystatin A toward papain was measured as described above. Protein concentrations were determined by the method of Lowry *et al.* (*34*) with bovine serum albumin as a standard.

RESULTS AND DISCUSSION

Cloning of Rat Cystatin A cDNA—Messenger RNA was prepared from newborn rat skin to construct a cDNA library, because it was reported that the cysteine proteinase inhibitor was detected in epidermal basal cells as polysome on fluorescence immunomicroscopy (*35*). The constructed cDNA library, which had about 10^5 independent clones, was screened by plaque hybridization using a ^{32}P-labeled oligonucleotide probe. Nine positive clones were obtained primarily, and further secondary screening and Southern hybridization gave only five positive clones. These inserts were subcloned into the pBluescript phagemid vector and then their nucleotide sequences were determined. All clones had a identical sequence and the length of the insert was 365 base pairs. The insert contained a partial coding sequence corresponding to Glu14-Phe103 (C-terminus) and then a putative poly(A) addition signal, AATAAA, followed by a poly(A)-tail (Fig. 1B). The deduced amino acid sequence completely coincided with that of rat epidermal cystatin A determined at the protein level (*16*). Other clones having a longer insert were searched for in the same library using the cDNA fragment as a probe, but no such clone was obtained.

Expression and Characterization of Recombinant Cystatin A—To prepare a recombinant cystatin A protein for the site-specific replacement study, an expression plasmid for the cDNA in *E. coli* was constructed. Since the cloned cDNA lacked a partial coding sequence corresponding to the amino-terminal region, initial Met-Ser13, of the inhibitor, a synthetic linker encoding the missing portion was designed by means of *E. coli* codon usage. The cDNA insert was introduced into an expression vector, pMK2, with the linker (Fig. 1A). A resultant expression plasmid for the rat cystatin A semi-synthetic gene, designated as pRC81, was used for the transformation of *E. coli* JM109. The transformant was cultured in LB/Amp overnight and then the whole cellular protein was subjected to SDS-polyacrylamide gel electrophoresis followed by Western blotting analysis (Fig. 2A). For *E. coli* cells transformed with pRC81, a single band corresponding to the molecular weight of cystatin A, 12 kDa, was immunostained (Fig. 2, A2, lane 4), although the expression level was too low to recognize the band on Coomassie brilliant blue staining (Fig. 2, A1, lane 4). Then, the gene was expressed without IPTG-induction in LB/Amp medium. However, there was no significant effect of the induction with various timings and concentrations of IPTG on the expression level (data not shown). To increase the expression level, we first tried to change the expression host-vector system from strain JM109-pMK2 (employing the *ori* region from pUC18, *tac* promoter) to BL21(DE3)-pET (having the *ori* region from pBR322, *T7* promoter). A coding region in pRC81 was excised and introduced into pET-15b with a linker for Met1-Thr5, resulting in a new plasmid, pET-CD. As *E. coli* BL21(DE3) was transformed with the plasmid and the gene expression was induced with 0.1 mM IPTG, an induced band appeared at a position corresponding to natural cystatin A on SDS-

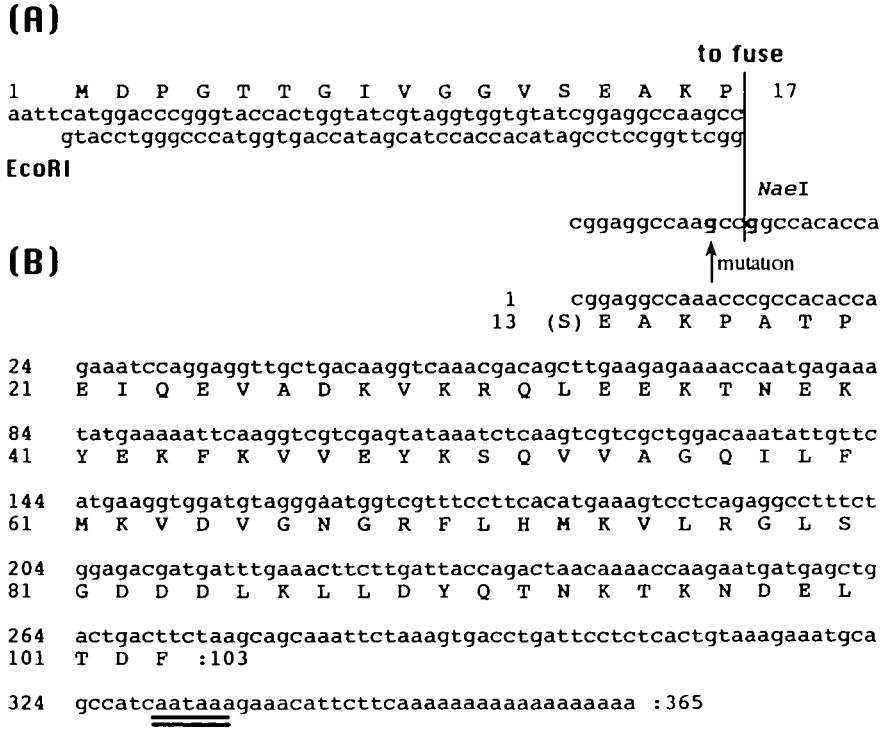

Fig. 1. **Nucleotide sequence and deduced amino acid sequence of rat cystatin A cDNA (panel B), and a cassette for expressing the cDNA in *E. coli* (panel A).** The deduced amino acid sequence coincided with that of rat cystatin A, corresponding to positions 14-103 (C-terminus). A putative poly A addition signal, AATAAA, is double-underlined (=). To fuse a cassette for expression of the cDNA, a *Nae*I restriction site was introduced at base positions 11-16 by site-specific mutagenesis. The base sequence of the cassette was designed using *E. coli* codon usage and includes a *Eco*RI site to link it with an expression vector, pMK2.

PAGE with CBB staining (Fig. 2B). The expression increased with this change in the host-vector system, however, the level was not enough, about 5% of the total cellular proteins, for further research. The potential formation of the secondary structure in mRNA around the initial Met codon was reported to influence its expression level (*36-38*). Prediction of the stem-loop formations around the initial Met of the pET-CD transcript (base numbers 1-110) showed that the segment from 61 to 71 around the initial ATG was complementary to the segment, 76-88, which corresponds to Gly 4-Ile 8. So we tried to change the degenerative third letter for these codons by random mutation rather than by designed alteration, because the effect on a break of a G-C pair by replacing G with any other, for example, could not be estimated and the various combinations of the substitutions could be examined simultaneously. Thus, random replacements of the degenerative third letters of codons for Gly 4, Thr 5, Thr 6, Gly 7, and/or Val 9 were made. Thirteen transformants were selected and assayed as to their cystatin A expression levels by SDS-PAGE (Fig. 3). These clones expressed the gene at a detectable level, as judged on CBB staining, and were found to have different combinations of the codons (Table I). The efficiency of expression of the gene for each clone was determined as the content of an expressed cystatin A in total cellular proteins on SDS-PAG E. Four of the 13 clones exhibited lower lesser levels compared to the original (pET-CD); however, the remaining 9 clones showed increased efficiency, especially pET-Cr:3-1, for which the level was increased about 3 times, 16.6%. To evaluate the relation between the expression level and the structural instability of the transcript, the minimum free energy for the formation of the secondary structure of the mRNA in several regions was calculated and compared (Fig. 4). These plots showed a correlation between the expression level and the structural instability of the 5′-terminal region (base numbers 1-110) of the transcript; a correlation coefficient of 0.7651 ($p<0.01$). Judging from the minimum free energy values obtained for other segments, 1-100, 1-120, 51-100, 51-110, and 51-120, there was a similar

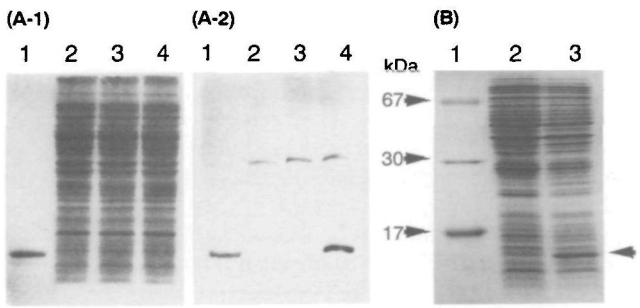

Fig. 2. **Expression of the semi-synthetic gene for rat cystatin A in *E. coli* JM109 (panel A) and BL21(DE3) (panel B).** (A) Total cellular proteins of *E. coli* with no vector (lane 2), and with the expression plasmids (lane 3, pMK2; lane 4, pRC81 including cystatin A cDNA) and isolated natural cystatin A (lane 1) were electrophoresed on SDS-polyacrylamide gel (15%), and then stained with CBB (panel A-1) or immunologically after Western blotting (panel A-2). (B) Expression of cystatin A by the pET system. The *E. coli* BL21(DE3) protein harboring an expression plasmid, pET-CD, with or without IPTG induction was compared (lanes 2 and 3, respectively). The induced protein corresponding to cystatin A is indicated by an arrow head on the right (lane 3). Lane 1; size markers.

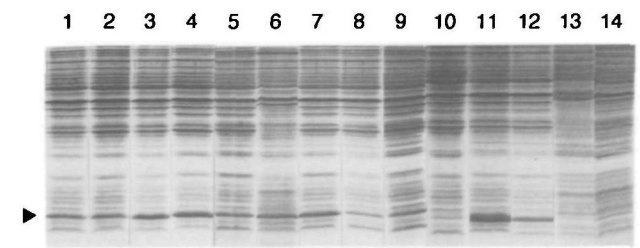

Fig. 3. **Selection of a high efficiency expression plasmid for rat cystatin A in the *E. coli* BL21(DE3)-pET system.** Total cellular proteins of *E. coli* cells harboring a series of expression plasmids with different codon usages were electrophoresed on a SDS-polyacrylamide gel (15%, CBB stained). The lane numbers correspond to the clone numbers in Table I. The arrowhead indicates the position of cystatin A. The expression efficiency (%) was estimated by image-analysis using software NIH-image.

TABLE I. **Nucleotide sequences of the N-terminal region of the rat cystatin A gene obtained by random replacement.**

Clone #	Name (pET-Cr:)	Amino acid sequence (4-9)					
		Gly	Thr	Thr	Gly	Ile	Val
		g g N	a c N	a c N	g g N	a t c	g t N
1	CD	- - T	- - C	- - T	- - T	- - -	- - A
2	3	- - T	- - T	- - G	- - G	- - -	- - T
3	4	- - T	- - T	- - C	- - T	- - -	- - T
4	7	- - T	- - T	- - A	- - T	- - -	- - C
5	10	- - T	- - G	- - G	- - C	- - -	- - C
6	14	- - T	- - T	- - A	- - T	- - -	- - C
7	15	- - G	- - G	- - A	- - T	- - -	- - T
8	16	- - C	- - T	- - C	- - T	- - -	- - G
9	19	- - A	- - A	- - G	- - T	- - -	- - T
10	3-5	- - T	- - C	- - T	- - G	- - -	- - T
11	3-1	- - C	- - T	- - A	- - T	- - -	- - T
12	5-1	- - T	- - G	- - A	- - G	- - -	- - G
13	5-2	- - T	- - C	- - G	- - T	- - -	- - G
14	5-3	- - G	- - G	- - G	- - G	- - -	- - G

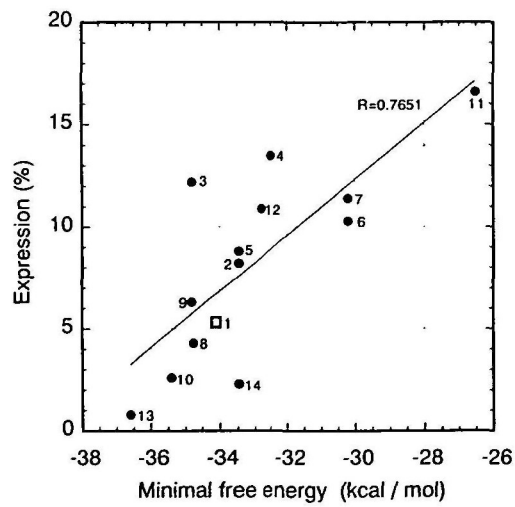

Fig. 4. **Relationship between the minimum free energy for secondary structure formation of the transcript and the cystatin A expression efficiency (ratio of cystatin A to total cellular protein, %).** The correlation coefficient (R) was 0.7651 ($p<0.01$). The numbers correspond to the clone numbers in Table I. The clone constructed originally, pET-CD, is shown by an open square.

tendency as to the correlation between them (data not shown).

The expressed cystatin A was purified from the soluble fraction of an *E. coli* cell lysate by successive chromatographies on DE-52, Sephadex G-75 and then Shodex DEAE-825 HPLC columns. The yield increased from 0.5 mg (pRC81) to about 8 mg purified protein (pET-Cr:3-1) per 1 liter culture. For the purified recombinant cystatin A, the N-terminal amino acid sequence was determined to be MDPGTTGIVG, which coincided with that of rat epidermal inhibitor except for the absence of an *N*-acetyl blocking group. The difference was also detected on comparison of API digested peptide maps on reverse phase HPLC, where only the retention time of the N-terminal fragment was shifted (data not shown). Other peptides derived from the recombinant protein exhibited the same retention times as those of the natural form, indicating that the amino acid sequences of the two cystatin As are identical. The far-UV CD spectra of the two cystatin As completely overlapped, suggesting there is no detectable difference in their secondary structures (data not shown). Their inhibitory activities toward papain were the same (Fig. 5), indicating that the N-acetyl group does not participate in the inhibitory action.

Site-Directed Mutagenesis—We previously found that the N-terminal region of rat cystatin A was essential for its inhibition of papain on comparison of the inhibitory activities of several N-terminally truncated forms (*17*), which were obtained by artificial processing during the preparation of a certain specific proteinase (*39*). Furthermore, series of engineered truncations (*29*) and substitutions (*40*) were introduced into the N-terminal region of human cystatin A, and it was revealed that the region may be responsible for maintaining the reactive sequence in the first loop, QVVAG, in an active conformation. Since some cystatins remain active even when they lack the N-terminal region including the conservative Gly, the main contribution to the interaction of the inhibitors with proteinases seems to be provided by the first loop. Comparison of the naturally occurring pentasequences showed that the first Q, the third V, and the last G were highly conserved, almost exclusively. The second position of the first loop is occupied by a hydrophilic Thr residue in cystatin S and SN, and similarly the fourth position tolerates Ser in chicken cystatin. In the case of human cystatin A, a first Gln was replaceable with Lys (*41*). Therefore, provided that the interaction of the inhibitors with target proteinases is mainly due to hydrophobic bonds of the first loop, the third Val seems to play a central role in the attraction between them. In fact, the third Val of cystatin B inserted deeply into the active cleft of papain is responsible for many interactions with Gln19-Gly23 and Trp177 of papain (*21*). Therefore, the third Val (V54) residue of rat cystatin A was substituted with a hydrophilic charged residue, Glu or Lys, in this study. But the two mutations, V54E and V54K, caused no large change in the inhibitory activity toward papain (Fig. 5), indicating that the position of Val54 needs not be occupied by a hydrophobic residue. Thus far, various replacements or deletions have been made in the first loop, however, their effects differ depending on the kinds of cystatin molecules and target proteinases used for the assessments. Therefore, in general, it can be concluded that the spatial arrangement of side chains of the first loop determines the complementarity against target proteinases. Also, the active conformation of the first loop of cystatin A might be controlled through interaction with the N-terminal segment. Now, we are performing structural analyses of an inactivated form of human cystatin A, G4V substitution mutant. The role of the N-terminal region in the control of the conformational change of the first and second loops will be reported elsewhere. Furthermore, to study the physiological function of rat cystatin A, the isolated cDNA fragment is useful for *in situ* hybridization or Northern hybridization analysis.

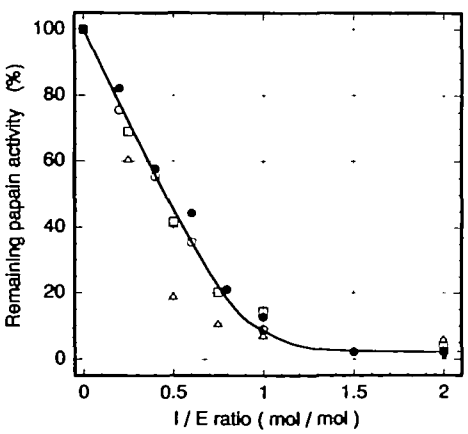

Fig. 5. **Assaying of the inhibitory activities of natural and recombinant cystatin A.** The remaining papain activity in the presence of various molecular inhibitor ratios was measured as described under "MATERIALS AND METHODS." ○, natural; ●, wild type; ■, V54E; ▲, V54K.

We wish to thank Dr. Kazunori Shibuya (Currently: Keio University, Japan) for the useful technical suggestion.

REFERENCES

1. Barrett, A.J., Fritz, H., Grubb, A., Isemura, S., Järvinen, M., Katunuma, N., Machleidt, W., Müller-Esterl, W., Sasaki, M., and Turk, V. (1986) Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem. J.* **236**, 312
2. Pennacchio, L.A., Lehesjoki, A.E., Stone, N.E., Willour, V.L., Virtaneva, K., Miao, J., D'Amato, E., Ramirez, L., Faham, M., Koskiniemi, M., Warrington, J.A., Norio, R., de la Chapelle, A., Cox, D.R., and Myers, R.M. (1996) Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). *Science* **271**, 1731-1734
3. Järvinen, M. (1976) Purification and properties of two protease inhibitors from rat skin inhibiting papain and other SH-proteases. *Acta Chem. Scand. B* **30**, 933-940
4. Järvinen, M., Räsänen, O., and Rinne, A. (1978) The low-molecular-weight SH-protease inhibitor in rat skin is epidermal. *J. Invest. Dermatol.* **71**, 119-121
5. Kominami, E., Bando, Y., Wakamatsu, N., and Katunuma, N. (1984) Different tissue distributions of two types of thiol proteinase inhibitors from rat liver and epidermis. *J. Biochem.* **96**, 1437-1442
6. Takahashi, M., Tezuka, T., and Katunuma, N. (1992) Phosphorylated cystatin α is a natural substrate of epidermal transglutaminase for formation of skin cornified envelope. *FEBS Lett.* **308**, 79-82
7. Takahashi, M., Tezuka, T., Kakegawa, H., and Katunuma, N. (1994) Linkage between phosphorylated cystatin α and filaggrin by epidermal transglutaminase as a model of cornified envelope

and inhibition of cathepsin L activity by cornified envelope and the conjugated cystatin α. *FEBS Lett.* **340**, 173-176
8. Steven, A.C. and Steinert, P.M. (1994) Protein composition of cornified cell envelopes of epidermal keratinocytes. *J. Cell Sci.* **107**, 693-700
9. Takahashi, H., Kinouchi, M., Wuepper, K.D., and Iizuka, H. (1997) Cloning of human keratolinin cDNA: keratolinin is identical with a cysteine proteinase inhibitor, cystatin A, and is regulated by Ca^{2+}, TPA, and cAMP. *J. Invest. Dermatol.* **108**, 843-847
10. Robinson, N.A., Lapic, S., Welter, J.F., and Eckert, R.L. (1997) S100A11, S100A10, annexin I, desmosomal proteins, small proline-rich proteins, plasminogen activator inhibitor-2, and involucrin are components of the cornified envelope of cultured human epidermal keratinocytes. *J. Biol. Chem.* **272**, 12035-12046
11. Takahashi, H., Asano, K., Kinouchi, M., Ishida-Yamamoto, A., Wuepper, K.D., and Iizuka, H. (1998) Structure and transcriptional regulation of the human cystatin A gene. The 12-*O*-tetradecanoylphorbol-13-acetate (TPA) responsive element-2 site (−272 to −278) on cystatin A gene is critical for TPA-dependent regulation. *J. Biol. Chem.* **273**, 17375-17380
12. Kartasova, T., Cornelissen, B.J., Belt, P., and van de Putte, P. (1987) Effects of UV, 4-NQO and TPA on gene expression in cultured human epidermal keratinocytes. *Nucleic Acids Res.* **15**, 5945-5962
13. Hawley-Nelson, P., Roop, D.R., Cheng, C.K., Krieg, T.M., and Yuspa, S.H. (1988) Molecular cloning of mouse epidermal cystatin A and detection of regulated expression in differentiation and tumorigenesis. *Mol. Carcinog.* **1**, 202-211
14. Laber, B., Krieglstein, K., Henschen, A., Kos, J., Turk, V., Huber, R., and Bode, W. (1989) The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS Lett.* **248**, 162-168
15. Bell, E.T., Featherstone, J.D., and Bell, J.E. (1989) Interaction of terbium and calcium with chicken cystatin. *Arch. Biochem. Biophys.* **271**, 359-365
16. Takio, K., Kominami, E., Bando, Y., Katunuma, N., and Titani, K. (1984) Amino acid sequence of rat epidermal thiol proteinase inhibitor. *Biochem. Biophys. Res. Commun.* **121**, 149-154
17. 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
18. 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
19. Engh, R.A., Dieckmann, T., Bode, W., Auerswald, E.A., Turk, V., Huber, R., and Oschkinat, H. (1993) Conformational variability of chicken cystatin. Comparison of structures determined by X-ray diffraction and NMR spectroscopy. *J. Mol. Biol.* **234**, 1060-1069
20. Dieckmann, T., Mitschang, L., Hofmann, M., Kos, J., Turk, V., Auerswald, E.A., Jaenicke, R., and Oschkinat, H. (1993) The structures of native phosphorylated chicken cystatin and of a recombinant unphosphorylated variant in solution. *J. Mol. Biol.* **234**, 1048-1059
21. Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, 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
22. Martin, J.R., Craven, C.J., Jerala, R., Kroon-Zitko, L., Zerovnik, E., Turk, V., and Waltho, J.P. (1995) The three-dimensional solution structure of human stefin A. *J. Mol. Biol.* **246**, 331-343
23. 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
24. Ekiel, I., Abrahamson, M., Fulton, D.B., Lindahl, P., Storer, A.C., Levadoux, W., Lafrance, M., Labelle, S., Pomerleau, Y., Groleau, D., LeSauteur, L., and Gehring, K. (1997) NMR structural studies of human cystatin C dimers and monomers. *J. Mol. Biol.* **271**, 266-277
25. 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
26. 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
27. Vaitukaitis, J., Robbins, J.B., Nieschlag, E., and Ross, G.T. (1971) A method for producing specific antisera with small doses of immunogen. *J. Clin. Endocrinol. Metab.* **33**, 988-991
28. Takeda, A., Kobayashi, S., and Samejima, T. (1983) Isolation and characterization of two thiol proteinase inhibitors of low molecular weight from newborn rat epidermis. *J. Biochem.* **94**, 811-820
29. Shibuya, K., Kaji, H., Itoh, T., Ohyama, Y., Tsujikami, A., Tate, S., Takeda, A., Kumagai, I., Hirao, I., Miura, K., 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
30. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Extraction, purification, and analysis of mRNA from eukaryotic cells in *Molecular Cloning (A Laboratory Manual)*, Cold Spring Harbor Laboratory, USA
31. Gubler, U. and Hoffman, B.J. (1983) A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269
32. Morinaga, Y., Franceschini, S.L., Inouye, S., and Inouye, M. (1984) Improvement of oligonucleotide-directed site-specific mutagenesis using double-stranded plasmid DNA. *Bio/Technology* **2**, 636-639
33. Kaji, H., Kumagai, I., Takeda, A., Miura, K., and Samejima, T. (1989) Studies on chemical synthesis of human cystatin A gene and its expression in *Escherichia coli*. *J. Biochem.* **105**, 143-147
34. Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265-275
35. Fukuyama, K., Ohtani, O., Hibino, T., and Epstein, W.L. (1982) Cellular localization of thiol-proteinase inhibitor in the epidermis of the newborn rat. *Cell Tissue Res.* **223**, 313-323
36. Iserentant, D. and Fiers, W. (1980) Secondary structure of mRNA and efficiency of translation initiation. **Gene 9**, 1-12
37. Tessier, L.H., Sondermeyer, P., Faure, T., Dreyer, D., Benavente, A., Villeval, D., Courtney, M., and Lecocq, J.P. (1984) The influence of mRNA primary and secondary structure on human IFN-γ gene expression in *E. coli*. *Nucleic Acids Res.* **12**, 7663-7675
38. Zabeau, M. and Stanley, K.K. (1982) Enhanced expression of cro-β-galactosidase fusion proteins under the control of the PR promoter of bacteriophage lambda. *EMBO J.* **1**, 1217-1224
39. Takeda, A., Nakamura, Y., Kaji, H., and Samejima, T. (1990) Evidence of protease activity producing NH_2-terminal truncated form of rat cystatin A in newborn rat epidermis. *Biol. Chem. Hoppe-Seyler* **371** Suppl, 233-238
40. Shibuya, K., Kaji, H., Ohyama, Y., Tate, S., Kainosho, M., Inagaki, F., and Samejima, T. (1995) Significance of the highly conserved Gly-4 residue inhuman cystatin A. *J. Biochem.* **118**, 635-642
41. Nikawa, T., Towatari, T., Ike, Y., and Katunuma, N. (1989) Studies on the reactive site of the cystatin superfamily using recombinant cystatin A mutants. Evidence that the QVVAG region is not essential for cysteine proteinase inhibitory activities. *FEBS Lett.* **255**, 309-314