Functional fusion expression of sunflower multicystatin in
E. coli and its comparison with a single domain cystatin

Ashraf Gholizadeh1 and B Baghban Kohnehrouz2

1Research Institute for Fundamental Sciences (RIFS), Department of Molecular Biology, 2Department of Plant Breeding and
Biotechnology, University of Tabriz, Tabriz, Iran

Received 05 July 2011; revised 08 November 2011

Identification of the molecular structure and novel biophysiological functions of plant cystatins or phytocystatins is of
great interest in the field of molecular biology. The important requirements for these are the efficient production,
purification and correctly folded forms of these proteins. We report here the cloning, easy expression and characterization of
a sunflower multicystatin (SMC) as a functional fusion protein in E. coli. For the first time, the amplified cystatin coding
region was expressed as a part of maltose-binding fusion protein using pMALc2X over-expression vector in TB1 strain of
E. coli without affecting the recombinant bacterial growth. In comparison to the previously prepared recombinant SMC
(rSMC), a high amount (~44 mg/L of bacterial cell culture) of purified fused SMC (fSMC) was obtained using single-step
purification method. fSMC strongly inhibited papain activity in vitro as compared to Celosia single-domain cystatin.
Purified fSMC may be used for basic biochemical, pharmacological or clinical studies without the cleavage of its fusion
parts.

**Keywords:** Cystatin, Expression, Fusion protein, Inhibitor, Sunflower multicystatin, Phytocystatin

The term cystatins refers to proteinaceous reversible inhibitors that specifically inhibit the enzymatic activity
of cysteine proteinases, such as papain and cathepsin H. The nucleotide sequences of their encoding genes have been characterized from different plants. Plant cystatins generally named as phytocystatins have wide
range of molecular mass ranging from 5-87 kD and contain three conserved regions that interact with
cysteine proteinases. They include ‘G’ residue at the N-terminus and ‘QxVxG’ and ‘W’ at the C-terminus. Moreover, they have a specific sequence [LVI]-[AGT]-[RKE]-[FY]-[AS]-[VI]-x- [EDQV]-[HYFQ]-N
at N-terminal α-helix region.

Various physiological and biological roles have been attributed for phytocystatins. However, to
increase the knowledge regarding their molecular structure and functions, the efficient production and
purification of their correctly folded three-dimensional structures is needed. These requirements are mostly
achieved by the heterologous protein expression as recombinant or fusion products in bacterial cells. However, thus far very limited information is available on the heterologous expression of phytocystatins.

In sunflower plant, a multicystatin cDNA sequence has been previously characterized from the seeds. A recombinant form of sunflower multicystatin (rSMC) has been reported to exhibit strong cysteine proteinase inhibitory activity in vitro. However, there has been no report on the structure or function of the fused protein products of SMC. As the fusion tags have been originally developed for enhancing the protein expression level, proper structural folding and stability in a native conformation, in this study, we have investigated the functional overexpression of SMC as a maltose-binding fusion product in E. coli cells and compared its efficiency with the rSMC and Celosia fused cystatin.

**Materials and Methods**

**Materials**

The seeds of *Helianthus annuus* L. (common sunflower) were provided by Dr B Baghban Kohnehrouz (Laboratory of Plant Genetic Engineering, Dept. Plant Breeding and Biotechnology). Trizol reagent used for total RNA extraction was purchased from Gibco BRL, USA (Cat. no. 15596-013). The mRNA purification kit was from QIAGEN, USA (Cat. No.70022). Chemicals used for the cDNA synthesis were provided from Promega (USA) cDNA synthesis kit (Cat. No. C4360). pGEM-T

---

*Corresponding author*
Fax: 98 (411) 3294113
E-mail: aghz_bioch@yahoo.co.in
Easy vector system I (Cat. no. A1360) was used in PCR product cloning. Restriction enzymes were purchased from Promega (Madison, WI, USA).

_E. coli_ strain TB1 and pMALc2X vector used for bacterial transformation, recombinant construction and protein expression studies were supplied in protein fusion and purification system kit (Cat. no. E8000S; NEW ENGLAND, Bio lab). Taq DNA polymerase, buffer, dNTPs and MgCl₂ used for PCR amplification were provided by CinnaGen. Fermentas DNA extraction kit (Cat. no. K0513) was used for purification of PCR product from the agarose gel. Restriction enzymes EcoRI and BamHI were purchased from CinnaGen Co. All the other chemicals were of analytical and molecular biology grades and purchased from Merck AG (Darmstadt, Germany) or Sigma (St. Louis, MO, USA).

**mRNA purification and cDNA synthesis**

Total cellular RNA was isolated from the sunflower seeds using trizol reagent. The integrity of the RNA was tested on 1% non-denaturing agarose gel using TBE (Tris/Borate/EDTA) running buffer. Poly (A⁺) RNA was purified from total RNA using mRNA purification kit. The double-stranded cDNA was synthesized according to the protocol of Promega cDNA synthesis kit.

**Specific cloning of sunflower cystatin**

The coding region of cystatin cDNA was amplified using polymerase chain reaction (PCR). PCR was carried out using sunflower seed cDNA population as template and a specific primer set (forward primer: 5'−GGCCGAATTCAGCTAAAAGATGTCACT−3’ and reverse primer: 5'−TGCTGGATCCTCATAAAGGTG CTTGA−3’) designed based on the previously reported sunflower seed cDNA sequence.[26]

For the directional cloning of PCR-amplified fragment in an _E. coli_ expression vector, EcoRI and BamHI restriction sites were included at the 5’ end of each primer. The reaction was carried out by mixing the following components in a 0.5 ml PCR tube: primers, 100 pmol (1 µl each); cDNA, 50 ng (1 µl); PCR reaction 10x buffer, 2.5 µl; 10 mM dNTP, 1 µl; H₂O, 17.5 µl and Taq polymerase, 1 µl. The reaction mixture was processed in thermocyclers (Techneh, Germany) under the following cycling program: denaturation at 94°C for 1 min, annealing at 58°C for 2 min, and extension at 72°C for 2 min. The PCR product was analyzed on a 1.2% agarose gel. The amplified fragment was then cloned in pGEM-T Easy vector system I and transformed to _E. coli_ strain DH5α. Transformants were grown in isopropyl β-D-thiogalactopyranoside (IPTG)/X-gal media and a single recombinant colony was selected and processed for plasmid extraction using the alkaline lysis method[27]. The isolated plasmid was digested with EcoRI restriction enzyme and separated on 0.8% agarose gel. Sequencing of the amplified fragment was done at Microsynth DNA sequencing center, Switzerland.

**Expression of cloned cystatin as fusion protein in _E. coli_**

The PCR amplification product after agarose gel purification step was digested with EcoRI and BamHI restriction enzymes, run on 1% agarose gel, extracted and purified from gel and ligated into the pMALc2X expression vector (Fig. 1). The ligation mixture was transferred into competent _E. coli_ TB1 cells. The transformed cells were plated on LB medium (supplemented with Amp and X-gal) at 37°C and a recombinant clone was selected for further gene expression studies[28].

**Extraction and purification of fused cystatin**

To extract the fusion protein, transformed cells were grown in 500 ml of rich broth medium containing glucose (2 g/l) and ampicillin (100 µg/ml). For the induction of fused protein expression, IPTG was added to a final concentration of 0.3 mM and incubated for 8 h. The cells were harvested by centrifugation at 4000 g for 10 min and the pellet was dissolved in 25 ml of extraction buffer containing 20 mM Tris-Cl, 200 mM NaCl, 1 mM EDTA, 1 mM azide and 10 mM BME (β-mecapto ethanol). Then the cells were frozen in the same buffer at -20°C overnight and sonicated in short pulses of 15 s. Sample was centrifuged at 10,000 g at 4°C for 20 min and the supernatant used as crude extract.

The fusion protein was purified from the crude extract by affinity column chromatography using a column packed with amylase resin specific for the maltose-binding protein (MBP), which was a part of fused protein. The fusion protein was eluted out from

---

**Fig. 1**—Cloned map of the SMC coding region on pMALc2X expression vector [EcoRI and BamHI digested SMC coding region was inserted with its correct reading frame, downstream of malE gene of _E. coli_ and resulted in an MBP (maltose-binding protein)-SMC fusion gene]
the column with column buffer (extraction buffer plus 10 mM maltose) and the eluted product was analyzed for its homogeneity by SDS-PAGE electrophoresis using 10% polyacrylamide gel.

**In vitro inhibitory test of fused protein**

The purified fusion protein was tested for cysteine proteinase inhibitory activity using papain (a cysteine proteinase from *Carica papaya*) and BANA (β-N-benzoyl-dL-arginine β-naphthlamide hydrochloride). A solution containing 0.1 ml of assay buffer (0.5 M sodium phosphate and 10 mM EDTA, pH 6.0), 0.1 ml of 50 mM 2-mercaptoethanol, 0.1 ml of papain solution (25 µg) and 0.2 ml of test solution containing 5, 10, 15, 20, 25, 30, 35, 40 and 50 µg purified protein respectively was incubated at 37°C for 10 min. The reaction was started by the addition of 0.2 ml of 1 mM BANA, incubated at 37°C for 20 min and then stopped with the addition of 1 ml of 2% (v/v) HCl/ethanol and 1 ml of 0.06% p-dimethyl-aminocinnamaldehyde/ethanol. The absorbance of reaction mixture was measured at 540 nm. A mixture containing lysate of non-recombinant *E. coli* cells was considered as control.

**Assessment of the recombinant bacterial growth**

To find out the effect of expressed fused products on the growth of recombinant bacteria, the growth of bacteria was assessed by measuring the optical density of the cultures at A$_{600}$. The assessments were carried out during 8 h incubation time with 1 h intervals.

**Results and Discussion**

**Cloning and sequence characterization of amplified cDNA**

SMC from seeds has been previously isolated using a degenerate primer sets and the screening of cDNA library. In the present study, SMC was amplified from the seed cDNA population using a specific primer pair based on the previously cloned sequence. The primers were designed by primer3 software as described under ‘Materials and Methods’. Separation of PCR end product on agarose gel and analysis of nucleotide sequence of amplified cDNA revealed that cDNA comprised an open-reading-frame of 819 nucleotides and was completely identical to the previously cloned SMC (Results not presented).

**Heterologous expression and purification of fused SMC**

Amplified SMC cDNA was expressed in *E. coli* cells as a part of maltose-binding fusion protein. Although SMC has been previously expressed as a recombinant protein (rSMC) in *E. coli* BL21-(DE3) using pET vector system, the purification and recovery of recombinant product from the expressing bacterial cells has been laborious and time-consuming process. Therefore, we used fusion tags for our study, as they facilitate recovery of expressed proteins by single-step protein purification process and also enhance proper folding and stability of fused proteins.

Expression of sunflower MBP-fused cystatin in TBI *E. coli* cells was induced by IPTG and analyzed in the soluble fraction of the induced bacterial cell culture. SDS-PAGE analysis revealed the presence of an expressed product with a molecular mass of about 72 kD which was in complete agreement with the calculated molecular mass for the expected fused product. The expected size consisted of MBP and the products of 273 amino acid residues of cDNA and six amino acid residues from the expression vector. The assessment of expression process of fused product by a time-course experiment revealed that fused protein was continuously expressed in TBI cells during induction time and there was a linear correlation between the amount of expression of fused product and time of induction (Fig. 2).
Comparison of the expression of fSMC in TB1 cells with rSMC in BL21-(DE3) cells has revealed that fSMC is expressed as a soluble protein, while rSMC predominantly occurs as insoluble protein in the form of inclusion bodies that needs to be processed with further solubilizing and refolding procedures.

In spite of difficulties in the purification procedure of earlier reported rSMC by Kouzuma group, in the present work, recovery of the MBP fused SMC was performed by single-step process using maltose-binding affinity chromatography and resulted in a high amount of purified product (purification yield of fSMC was about 44 mg/L of bacterial cell culture), as compared to rSMC. Such high amount of pure fused protein obtained after a single-step of purification might validate the simplicity of the pMALc2X expression and purification system for production of SMC in bacterial cells, as compared to the rSMC expression and purification system.

**In vitro inhibitory activity of fused SMC**

We previously reported that *Celosia* cystatin (celostatin) can be expressed as an active maltose-binding fusion protein in *E. coli* cells. In the present work, we examined the activity of the expressed product of fSMC and compared it with the *Celosia* fused cystatin.

In the first step, our bioinformatics analysis of the primary structures of sunflower and *Celosia* cysteine proteinase inhibitors at NCBI site revealed the presence of three putative active sites in the SMC, as compared to one active site in *Celosia* cystatin (Fig. 3). Based on the primary structures, a putative activity ratio of about 3:1 was predicted for sunflower and *Celosia* cystatin molecules.

In the second step, similar to celostatin, the inhibitory activity of the expressed fSMC was examined against papain. The results showed that there was an inverse relationship between the concentration of the fusion product and papain activity of the test solution (Fig. 3, shown as open squares). The papain activity decreased with the increasing concentrations of fused protein. This result confirmed that fSMC was expressed and purified in correct conformation in bacterial cells, thus can be employed for the further biochemical, pharmacological or clinical studies. The present study also confirmed that fSMC can be functionally used without the cleavage of the fusion parts.

![Fig. 3—Comparison of sunflower and *Celosia* cystatin molecules](image-url)
Comparison of the inhibitory activity of MBP-fused sunflower against *Celosia* fused cystatin confirmed that the activity of SMC was significantly higher than that of *Celosia* cystatin. Our estimations revealed the predicted activity ratio of about 3:1 at the concentration ranging from 15-20 µg/ml for purified SMC and *Celosia* cystatins (Fig. 3, data not presented). However, further studies are needed to demonstrate how each domain in SMC molecule contributes in the overall inhibitory activity. A number of physico-chemical characteristics of sunflower and *Celosia* cystatin molecules were computed using Expasy proteomic tools at Proparam (Fig. 3). The computed instability data showed that *Celosia* cystatin was more unstable than the SMC.

**Effect of SMC expression on bacterial growth**

The growth of bacterial cells carrying fSMC was evaluated by measuring the absorbance of the cultures at A600. The results showed that bacterial growth was not significantly affected by the expression of fSMC molecule in TB1 cells under induced condition using IPTG. On the other hand, the same growth patterns were observed under induced and non-induced conditions (data not presented). This result indicated that the expression of fSMC protein had no inhibitory effect on *E. coli* cells.

In conclusion, TB1 cells can be used as suitable bacterial factory for the high level expression of enzymatically active SMC using maltose-binding fusion construct.

**Acknowledgments**

The authors are thankful to the Research Institute for Fundamental Sciences (RIFS), University of Tabriz, for financial supporting of this work.

**References**