Biochimica et Biophysica Acta 1804 (2010) 1413–1423

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Functional expression of horsegram (*Dolichos biflorus*) Bowman–Birk inhibitor and its self-association

Deepa G. Muricken, Lalitha R. Gowda *

Department of Protein Chemistry and Technology, Central Food Technological Research Institute, Council of Scientific and Industrial Research (CSIR) Mysore, 570-020, India

ARTICLE INFO

Article history:
Received 10 September 2009
Received in revised form 12 February 2010
Accepted 24 February 2010
Available online 20 March 2010

Keywords:
Recombinant horsegram inhibitor
Cloning
Site-directed mutagenesis
Monomer–dimer interaction
C-terminal Asp

ABSTRACT

Horsegram (*Dolichos biflorus*), a protein-rich leguminous pulse, native to Southeast Asia and tropical Africa, contains multiple forms of Bowman–Birk inhibitors. The major Bowman–Birk inhibitor from horsegram (HGI-III) was cloned and functionally expressed in *Escherichia coli* (rHGI), which moved as a dimer in solution similar to the natural inhibitor. The biochemical characterization of rHGI also points to its close resemblance with HGI-III not only in its structure but also in its inhibitory characteristics. To explore the electrostatic interactions involved in the dimerization, a site-directed mutagenesis approach was used. The role of reactive site residue K^{24} and the C-terminal Asp in the structure and stability of the dimer was accomplished by mutating K^{24} and $D^{75/76}$. The mutants produced in this study confirm that the self-association of HGI-III is indeed due to the electrostatic interaction between K^{24} of one monomer and $D^{75/76}$ of the second monomer, in agreement with our previous data. The functional expression of a Bowman–Birk inhibitor minus a fusion tag serves as a platform to study the structural and functional effects of the special pattern of seven conserved disulphide bridges.

1. Introduction

Proteases, although essentially indispensable to the maintenance and survival of their host organisms, can be potentially damaging when overexpressed or present in higher concentrations, and their activities need to be regulated. Proteinase inhibitors therefore play a vital role in regulating proteases. Many proteinaceous proteinase inhibitors have been isolated from plant storage organs [1,2]. The large amount of reserve protein in legume seeds are noted for their proteinase inhibitor content, the synthesis being elicited by mechanical wounding or insect or pathogen attack [2]. These proteinaceous inhibitors play a key role in plant arsenal against insect herbivory and combating proteinases of pests and pathogens [3,4].

Two distinct families, the Bowman–Birk inhibitors (BBIs) and Kunitz-type inhibitors, of legumes are among the most well studied. This prominence is due to their specific inhibition of trypsin-like family of serine proteases [1,5–8]. These inhibitors bind to their cognate proteinase according to a common substrate canonical mechanism, similar to that of a productively bound substrate [5]. The BBIs characteristically are single polypeptides with molecular masses in the range of 6–9 kDa and comprise of a binary arrangement of two subdomains with a conserved array of seven disulphide linkages that play a vital role in stabilizing the kinetically independent configuration of the reactive site on the outer most exposed loop [9,10]. Despite the extensive studies on BBIs, only a few three-dimensional structures have been solved either by X-ray or by NMR. These include PI-II from tracy bean [11], A-II from peanut [12], and soybean BBI [13–15] pea seed PsTI-IVb [16], snail medic seeds MSTI [17], seeds of *Vigna ungulata*, Vu-BBI, and BTCI [18,19]. Renewed interest in BBIs arises from their ability to suppress carcinogenesis both in vivo and in vitro model systems [20–22]. Consequently, a BBI concentrate (BBIC) achieved "investigational new drug status" as an anticarcinogenic agent by the FDA in 1998. Besides the anticarcinogenic effects, BBIs also inhibit inflammation-mediating proteases potentiating an anti-inflammatory role [23]. BBIC has recently been used in treating ulcerative colitis [24] and multiple sclerosis [25]. Considering the positive aspects of BBIC, incorporation into commercial foods has been attempted [26].

Horsegram (*Dolichos biflorus*) is a pulse crop native to Southeast Asia and tropical Africa. Four BBIs from horsegram have been isolated and characterized. The major BBI of horsegram (HGI-III) is a 76-amino acid single-chain polypeptide with two independent inhibitory domains directed toward trypsin and chymotrypsin [27]. The complete primary structure of HGI-III [28] and the role of disulphide bonds in maintaining the structural integrity of HGI-III have been established [29,30]. HGI-III is a group II BBI characterized by the Pro-Ala sequence near the second reactive site. Three sequential epitopes

Abbreviations: APNE, *N*-acetyl-DL-phenylalanine-β-naphthyl ester; BAPNA, *N*-α-, benzoyl-DL-arginine-*p*-nitroanilide; BTPNA, *N*-benzoyl-L-tyrosine-*p*-nitroanilide; NAPNA, *N*-succinyl-Ala-Ala-Ala-*p*-nitroanilide; BBI, Bowman–Birk inhibitor; CIU, chymotrypsin inhibitory unit; DMSO, dimethyl sulfoxide; HGI, horsegram inhibitor; rHGI, recombinant horsegram inhibitor; TPCK, L-1-tosylamido-2-phenylalanine chloromethyl ketone; TIU, trypsin inhibitory unit; T, total acrylamide concentration; C, degree of cross-linking

* Corresponding author. Tel.: +91 821 2515331; fax: +91 821 2517233.
E-mail addresses: lrg@cftri.res.in, lrgowda@yahoo.com (L.R. Gowda).

1570-9639/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2010.02.012

of HGI-III have been mapped, of which one includes the chymotrypsin inhibitory site [31]. HGI-III, though a single polypeptide of ~8600 Da, like several other BBIs, undergoes self-association and exists as a dimer in solution [32]. Several BBIs undergo self-association to form homodimers or trimers or more complex oligomers in solution [33]. In direct contrast, the three inhibitors of germinated horsegram seeds (HGGIs) are single polypeptides of 6500–7200 Da that exist as monomers and exhibit no such self-association [34]. Chemical modifications coupled with a comparative evaluation of several BBI protein sequences and homology modeling indicated that the dimerization was characterized by a crucial electrostatic interaction that comprised the C-terminal tail of subdomain II [32].

Despite the vast potential application of BBIC as therapeutics, reports on efficient production systems are currently limited to the expression of BBIs as fusion proteins. The gene-coding soybean (*Glycine max*) BBI was chemically synthesized and expressed in *Escherichia coli* as a fusion protein, which, on CNBr cleavage, required refolding [35]. Prokaryotic expression of rice BBI showed that the fusion protein exhibited only trypsin inhibitory activity [36]. Soybean BBI expressed as a fusion protein in *Bacillus subtilis* also required activation [37]. Pea seed inhibitor was expressed using a protein fusion expression where the inhibitor was linked to *Aspergillus niger* glucoamylase protein [38]. The functional expression of a recombinant BBI is therefore required not only for therapeutic applications but also as a promising model to study mechanism, specifically any molecular changes that might improve its efficacy. In the present investigation, we report the cloning and functional expression of HGI-III the major Bowman–Birk inhibitor of horsegram. Further, by site-directed mutagenesis, we validate the involvement of the C-terminal tail in a crucial interaction that is vital for self-association.

2. Materials

2.1. Plant material

Horsegram seeds (*D. biflorus*) were procured from the local market, which served as starting material.

2.2. Bacterial strains and vectors

The bacterial strain *E. coli* DH5α and pRSETC vector were from Invitrogen Corporation, Carlsbad, CA. *E. coli* BL21 (DE3) pLysS and Origami (DE3)pLysS and the vector pET20b were obtained from Novagen, Merck Specialities Private Ltd, Mumbai, India.

2.3. Chemicals

Tryptone type 1, yeast extract, bacto-agar, ampicillin sodium salt, and isopropyl thiogalacto pyranoside (IPTG) were from Hi-Media Laboratories, Mumbai, India. Genomic DNA extraction kit, gel extraction kit, and *Taq* DNA polymerase were from Qiagen, GmbH, Hilden, Germany. *Pfu* DNA polymerase for gene amplification was obtained from Fermentas Life Sciences GmbH, Hilden, Germany. Restriction enzymes for gene manipulation and expression were purchased from New England Biolabs, Beverly, MA, USA. Bovine serum albumin (BSA), bovine pancreatic trypsin (2× crystallized, type III, EC 3.4.21.4), L-1-tosylamido-2-phenylalanine chloromethyl ketone (TPCK)-treated trypsin, bovine pancreatic α-chymotrypsin (3× crystallized, type II, EC 3.4.21.1), porcine pancreatic elastase (EC 3.4.21.36), *N*-α-benzoyl-DL-arginine-*p*-nitroanilide (BAPNA), *N*-benzoyl-L-tyrosine-*p*-nitroanilide (BTPNA), *N*-succinyl-Ala-Ala-Ala-*p*-nitroanilide (NAPNA), CNBr-activated Sepharose® 4B, 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS), *N*-acetyl-DL-phenylalanine-β-naphthyl ester (APNE), oxidized glutathione, and imidazole were procured from Sigma-Aldrich, St. Louis, MO, USA. Polyvinylidine difluoride membrane (PVDF; 0.45 μm) was from Pierce, USA. Ni-Sepharose™ 6 Fast Flow was purchased from Amersham Biosciences AB Uppsala, Sweden. Cellulose acetate (5000 Da cutoff) was obtained from Millipore Systems, Bedford, MA, USA. The oligonucleotides were synthesized by Sigma-Genosys, Sigma-Aldrich Chemical Private Limited, Bangalore, India. Molecular weight markers for SDS–PAGE were from Bangalore Genei, Bangalore, India. All other chemicals used were of highest purity.

3. Methods

3.1. Isolation of genomic DNA

Horsegram seeds were milled to a fine powder and defatted with CCl_4 (1:5 wt./vol.) for 14–16 h at 25 ± 2 C. Genomic DNA was extracted from 0.2 g of defatted seed powder using a silica gel-based membrane-type kit (DNeasyPlant, Qiagen GmbH, Hilden, Germany) following the manufacturer's protocol. The DNA concentration was determined by spectrophotometry (UV-1601, Shimadzu, Japan) after dilution. The DNA concentration and the impurity factor (A_{260}/A_{280} ratio) were recorded. This genomic DNA was used in subsequent PCR analyses.

3.2. Cloning and construction of recombinant plasmids for total prokaryotic expression of HGI-III

Genomic DNA isolated from horsegram seeds were used as the template for PCR amplification of HGI-III coding sequence. Primer sequences designed using the published sequences of HGI-III [28] (GenBank Accession No. AY049042), were used to obtain the open reading frame of 228 bp. The PCR primers corresponded to the first 8 N-terminal amino acids of HGI-III and last 8 C-terminal amino acids of the mature coding sequence. The primer sequences were HGI-F (5′ GATCATCATCAGTCAACTGATGAG3′) and HGI-R (5′ATCATCATGTGAA-GATTTGCAAGG3′). Touchdown PCR conditions used to amplify a 228 bp fragment were 94 °C, 15 s; 54–44 °C, 30 s; 72 °C, 20 s; for 10 cycles followed by 94 °C, 15 s; 44 °C, 30 s; 72 °C, 20 s; for 10 cycles in a GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA, USA). The amplified PCR product was separated on agarose gel electrophoresis. The PCR product was purified using a PCR product purification kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's instructions and subjected to direct dideoxy sequencing on an automated DNA sequencer (ABI 310, Applied Biosystems Foster City, CA, USA). Primers were designed to introduce *Nhe*I and *Nde*I sites at the 5′ end (HGI-Nde-F; Table 1) and *BamH*I (HGI-BamH1-R; Table 1) at the 3′ end to facilitate cloning of the 228 bp into pRSETC vector (Invitrogen Corporation, Carlsbad, CA, USA). Using these two oligonucleotide primers and the 228-bp amplicon as template DNA, a PCR product containing the entire HGI-III coding region with *Nhe*I and *Nde*I site at the 5′ end and *BamH*I site at the 3′ end was obtained. The PCR product was blunt end cloned into the *Pvu*II site of pRSETC vector to generate a vector named pRSET-rHGI (3157 bp) (Fig. 1). Chemically

Table 1
Oligonucleotide sequences used for cloning and site-directed mutagenesis of rHGI.

Primer	Oligonucleotide sequence
K24A-F	5′TGCGCATGCACAGCGTCAATCCCTCCTCAATGC3
K24A-R	5′GCATTGAGGAGGGATTGACGCTGTGCATGCGCA3′
D75A-F	5′ AAATCTTCACATGCTGATCTCGAGCACCAC3′
D75A-R	3′TTTAGAAGTGTACGACTAGAGCTCGTGGTG5
HGI-*Nde*-F	5′CTAGCTAGCC*ATATG*GATCATCATCAGTCA3′
HGI-*Bam*H1-R	5′CGC*GGATCC*TTAATCATCATGTGAAG3′
HGI-*Xho*-R	5′CGCGGATCCTTAC*TCGAG*ATCATCATGTGAAG3′
Δ76-*Xho*-R	5′*CTCGAG*TTAATCATGTGAAGATTTGCAAGGTGC3′

The underlined residues in the table indicate the region of point mutation. The restriction sites used are italicized.

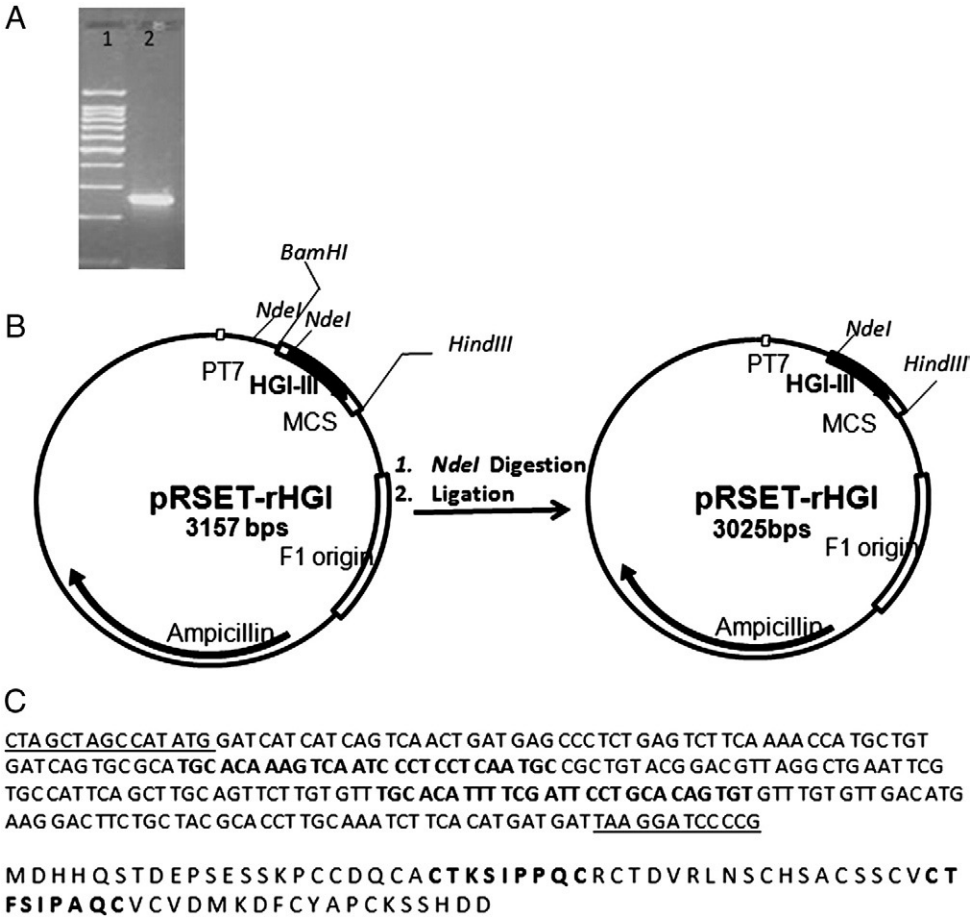


Fig. 1. (A) Agarose gel electrophoresis to show the 230-bp amplicon from genomic DNA. Lane 1, 100 bp DNA ladder; lane 2, PCR product (~230 bp). (B) Schematic representation of the expression cassette in pRSETC vector to produce rHGI protein. (C) The nucleotide sequence of genomic DNA in the open reading frame of rHGI and translated amino acid sequence of rHGI. The restriction sites *Nde*1 and *Bam*H1 are underlined in the nucleotide sequence. Trypsin and chymotrypsin inhibitory domains in the amino acid sequence are shown in bold.

competent *E. coli* DH5α cells were transformed with pRSET-rHGI. Positive clones harboring the insert were verified by gel shift assays and *BamH*I digestion. The sequence of the inserted fragment was confirmed by dideoxy sequencing using the Big Dye^R terminator v3.1 cycle sequencing kit (Applied Biosystems, USA), products purified using Dye Ex™ 2.0 spin kit (Qiagen GmbH, Hilden, Germany) and sequenced using an ABI 310 DNA Genetic Analyser (Applied Biosystems, Foster City, USA). The codons for the 44 amino acids in excess at the N-terminal end of the vector pRSET-rHGI (3157 bp) were removed by *Nde*I digestion, allowed to self-ligate generating the vector pRSET-rHGI (3025 bp) (Fig. 1), which was transformed into chemically competent *E. coli* DH5α. The transformants were identified by gel shift assay and insert release. Plasmid DNA was purified using the alkaline lysis method and DNA sequence determined.

3.3. Expression of recombinant HGI (rHGI)

E. coli BL21 (DE3) pLysS transformed with the expression vector pRSET-rHGI, was grown overnight at 37 °C in Luria–Bertani medium (10 mL) supplemented with ampicillin (100 μg/mL). Ten milliliters of the cells was diluted 25-fold into 250 mL of 2YT medium supplemented with ampicillin (100 μg/mL) and incubated with shaking at 37 °C until the optical density (OD) at 600 nm reached 1.75. IPTG was added to a final concentration of 0.3 mM, and the culture was further incubated at 37 °C for 4 h. The cells were harvested at 8000 rpm for 15 min at 4 °C and then lysed in 0.1 M Tris–HCl pH 8.2 by ultrasonication (Vibracell™; Sonics and Materials, Inc., New Town, CT, USA). Both the supernatant and the pellet were evaluated for trypsin inhibitory activity. The supernatant containing rHGI was further purified.

3.4. One-step trypsin Sepharose affinity purification of recombinant HGI

The supernatant (cell free extract) was applied to a trypsin–Sepharose column (12 × 3.4 cm) preequilibrated with 0.1 M Tris–HCl pH 8.2 containing 0.1 M NaCl at a flow rate of 10 mL/h. The column was thoroughly washed with the same buffer until the A_{230} was zero. The bound rHGI was eluted with 0.2 M glycine–HCl pH 3.0 containing 0.1 M NaCl at a flow rate of 30 mL/h. Two-milliliter fractions were collected and assayed for trypsin inhibitory activity. The pH of the pooled trypsin inhibitor fraction was adjusted to 7.5 and dialyzed against water and lyophilized.

3.5. Trypsin, chymotrypsin, and elastase inhibitory assay

The amidase activity of trypsin and its inhibition were assayed using the chromogenic substrate BAPNA at pH 8.2 in 0.05 M Tris–HCl containing 0.02 M $CaCl_2$ at 37 °C according to the method of Kakade et al. [39]. The amidase activity of chymotrypsin and its inhibition were assayed using BTPNA at pH 7.8 in 0.08 M Tris–HCl containing 20% DMSO and 0.02 M $CaCl_2$ at 37 °C. One unit of trypsin/chymotrypsin enzyme activity is defined as the increase in the absorbance of 0.01 at 410 nm under the assay conditions. One inhibitory unit is defined as the amount of inhibitor that reduces the enzyme activity by one unit. Elastase inhibition was assayed in a similar way at pH 8.2 using NAPNA as substrate.

3.6. Protein estimation

Protein content was measured by the dye binding method of Bradford [40]. BSA was used as the standard.

3.7. Polyacrylamide gel electrophoresis (PAGE)

Native PAGE (10% T, 2.7% C) of the purified rHGI was carried out according to the method of Laemmli [41]. Gelatin-embedded PAGE was performed by incorporating gelatin 0.5% (wt./vol.) in the gel [42]. Following electrophoresis, the gel was washed with distilled water and incubated at 37 °C in 0.1 M Tris–HCl buffer, pH 8.0, containing trypsin (40 μg/mL) or at pH 7.8 containing chymotrypsin (40 μg/mL). After gelatin hydrolysis, the gel was washed and stained with Coomassie brilliant blue R-250 and destained. The presence of rHGI was detected as a dark blue band with a clear background due to the complex of the unhydrolyzed gelatin and stain. For APNE staining, the gel was incubated in either 0.1 M sodium phosphate buffer, pH 8.0, containing trypsin (5 mg in 20 mL) or 0.1 M Tris–HCl buffer, pH 7.8, containing chymotrypsin (40 μg/mL) for 1 h, and visualized with APNE (4 mg/mL in DMSO), and *o*-dianisidine tetrazotised (7 mg/10 mL of 0.1 M sodium phosphate buffer, pH 7.4) for 15 min. The appearance of a clear transparent band against a pink background indicates the presence of the inhibitor.

3.8. Molecular weight (mass) determination

The molecular weight of purified rHGI and mutant HGIs were determined by SDS–PAGE following the published procedure [41]. All the SDS–PAGE experiments were carried out under reducing conditions at pH 8.3. The purified proteins in sample buffer were boiled for 5 min and separated by SDS–PAGE (15% T, 2.7% C) at pH 8.8. The separated proteins were visualized with either 0.1% Coomassie brilliant blue or by silver staining. The gel was stained with 0.25% Coomassie brilliant blue R-250 and destained with 10% (vol./vol.) acetic acid. Ovalbumin (M_r 45,000), carbonic anhydrase (M_r 29,000), soybean trypsin inhibitor (M_r 20,000), lysozyme (M_r 14,500), and aprotinin (M_r 6500) were used as marker proteins.

Electrospray ionization–mass spectrometry (ESI-MS) was carried out in the ionisation mode using an Acquity SYNAPT HRMS (Waters Associate, Miliford, USA). The RP–HPLC-purified protein was suspended in 50% CH_3CN containing 0.1% HCOOH and was directly infused, and the spectra were analyzed and deconvoluted using Max Ent1.

3.9. Size-exclusion chromatography

Size-exclusion measurements were performed using a BIOSEP-SEC-S 3000 (300×8 mm, exclusion limit: 700 kDa for globular proteins) column on a Waters Associate HPLC equipped with a binary gradient pumping system and Waters Model 1296 photodiode array detector set at 230 nm. The column was preequilibrated with the corresponding buffers at a flow rate of 1 mL/min before sample loading. The column was calibrated using a mixture of standard proteins, bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), cytochrome *c* (14.4 kDa), and aprotinin (6.4 kDa).

3.10. N-terminal sequence analysis

The purified protein after SDS–PAGE was transferred to a PVDF membrane in 0.01 M CAPS–10% methanol buffer (pH 11) and stained with Coomassie brilliant blue R-250. The bands corresponding to rHGI were excised, washed with methanol, and loaded directly to the Applied Biosystems 477A automated gas phase sequencer for N-terminal sequencing by Edman degradation. β-Lactoglobulin was sequenced as a standard to validate the performance of the instrument.

3.11. Construction of point mutants of rHGI

Table 1 lists the primer pairs used for site directed mutagenesis. Mutants were made using the Quik Change PCR-based mutagenesis method [43] and pRSET-rHGI as template, with one modification. *Pfu* DNA polymerase was used for amplification. Methylated (parental) DNA was degraded using *Dpn*I and nonmutant template digestion evaluated by agarose gel electrophoresis. The mutated DNA was transformed in chemically competent *E. coli* DH5α. Plasmid DNA was isolated by alkaline lysis method and sequenced according to the dideoxy chain termination method to check for random PCR errors. The mutants were expressed in *E. coli* strain BL21 (DE3) pLysS and purified on Sephadex G-100 by size exclusion.

3.12. Construction of expression plasmids and point mutants of HGI-III as C-terminal histidine-tagged proteins

A C-terminal $(His)_6$-tagged expression of HGI-III was used to abet in protein purification. The HGI-III coding region was PCR-amplified using the primer pair HGI-F (5′GATCATCATCAGTCAACTGATGAG3′) and HGI-Xho-R (Table 1) from pRSET-rHGI digested with *Xho*1 and ligated into pET-20b digested with *Eco*RV and *Xho*1 restriction enzymes. The expressed protein from the recombinant clone (pET-20b-HGI) would have, in addition to the 76 residues of HGI-III, two amino acids (LE) plus the 6 residue His tag ($(His)_6$) at the C-terminus and two amino acids (MD) at the N-terminus. Using pET-20b-HGI as the template and the corresponding set of mutant primers (Table 1), the mutant clones pET-20bHGI K24A and pET-20bHGI D75A were constructed by the PCR-based Quik Change method as described above. After cloning into *E. coli* DH5α cells and plasmid DNA isolation, the mutants sequence was validated by the dideoxy chain termination method to check for random PCR errors. These clones were used for the expression of His-tagged rHGI and its mutants in *E. coli* strain BL21 (DE3) pLysS cells.

3.13. Construction of Δ76 mutant of rHGI

Using pET-20b-HGI DNA and the primer pair HGI-F and Δ76-Xho-R, the deletion mutant pET 20bHGIΔ76 was constructed as described above for pET-20b-HGI and expressed in *E. coli* strain BL21 (DE3) pLysS. The expressed protein was purified by Sephadex G-100 size exclusion chromatography followed by RP–HPLC using a C-18 column.

3.14. One-step purification of the $(His)_6$-tagged fusion proteins

The cell-free extract was applied to a Ni^{2+}-NTA Sepharose column preequilibrated with 0.02 M Tris–HCl containing 0.5 M NaCl and 0.03 M imidazole pH 7.4. The column was washed with the same buffer until A_{280} returned to baseline. Unbound proteins were eluted in this buffer. The bound $(His)_6$-tagged fusion protein was eluted with the above buffer containing 0.5 M imidazole and fractions checked for trypsin inhibitory activity. The fractions containing the fusion protein were collected and dialyzed overnight against water at 4 °C.

3.15. Refolding of purified rHGI in solution

Refolding of the purified rHGI and mutants in solution were carried out as reported for CNBr cleaved recombinant soybean BBI [35]. Purified protein (15 mg) was unfolded in 0.24 mL of unfolding buffer (2 M Tris–HCl, pH 8.0, 0.2% EDTA, 6 M GuHCl, and 2 M β-mercaptoethanol) for 18 h at 37 °C. Refolding was performed by adding the unfolded inhibitor dropwise into 240 mL of refolding buffer (0.08 M Tris–HCl, pH 8.0, 0.1 mM EDTA, and 0.2 mM oxidized

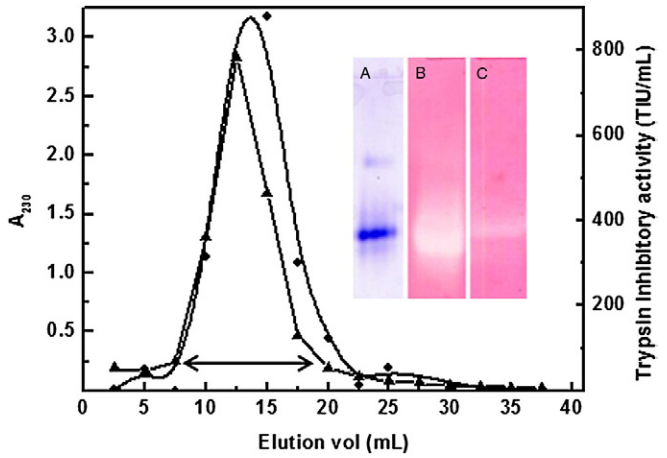


Fig. 2. Trypsin Sepharose affinity chromatography profile of rHGI. The *E. coli* cell lysate was loaded on a trypsin Sepharose column equilibrated with 0.1 M Tris–HCl buffer (pH 8.2) containing 0.1 M NaCl at a flow rate 10 mL/h. (—●—) TIU/mL, (—▲—) A_{230}. The bound protein was eluted by altering the pH with 0.2 M glycine–HCl, pH 3.0, containing 0.1 M NaCl at a flow rate of 30 mL/h. The arrow shows the active inhibitor fractions that were pooled. Inset: (A) Gelatin-embedded native PAGE profile of rHGI showing trypsin inhibitory activity and native PAGE (10% T, 2.7% C) of rHGI at pH 8.8. The gels were stained for (B) trypsin inhibitory activity and (C) chymotrypsin inhibitory activity with APNE.

glutathione with vigorous stirring, for 18 h at 37 °C). The refolded sample was concentrated by ultrafiltration with a 5000-Da cutoff cellulose acetate flat membrane on a Mini Labscale™ TFF system (Millipore Systems, Bedford, MA, USA), dialyzed and lyophilized.

3.16. Inhibition studies and stoichiometry

The effect of varying substrate concentration (BAPNA, BTPNA, and NAPNA) with bovine trypsin, chymotrypsin, and elastase in the presence of fixed concentrations of rHGI was studied. The modes of inhibition and dissociation constants were evaluated from the double reciprocal [44] and Dixon plots [45] of the data, respectively. The stoichiometry was studied by titrating rHGI with a constant optimum concentration of bovine trypsin under conditions of the assay, and residual activity was measured. Extrapolation of the plot of residual trypsin activity against the inhibitor/enzyme concentration ratio provided the stoichiometry of binding.

3.17. Stability studies

The purified recombinant inhibitor was dissolved in 0.1 M Tris–HCl buffer and incubated at 90 ± 1 °C in a constant temperature water bath. Aliquots were removed at regular time intervals, immediately cooled on ice, and assayed for residual inhibitor activity as described earlier. The purified rHGI was dissolved in 0.1 M buffers of pH 3, 5, 7, and 8.2 and incubated for 1 h at 37 °C. The residual trypsin inhibitor activity was assayed with BAPNA as described earlier.

3.18. Disulphide bond assay using 2-nitro-5-thiosulfobenzoate (NTSB)

The disulphide content of the rHGI was evaluated using NTSB as described earlier [46]. For the disulphide assay 1.27×10^{-8} mol of rHGI in 0.05 mL of buffer was added to 3.0 mL of NTSB assay solution and incubated at 37 °C. The change in absorbance at 412 nm due to the formation of 2-nitro-5 thiobenzoic acid (NTB) was observed for 60 min. Ribonuclease was used as the reference standard. The

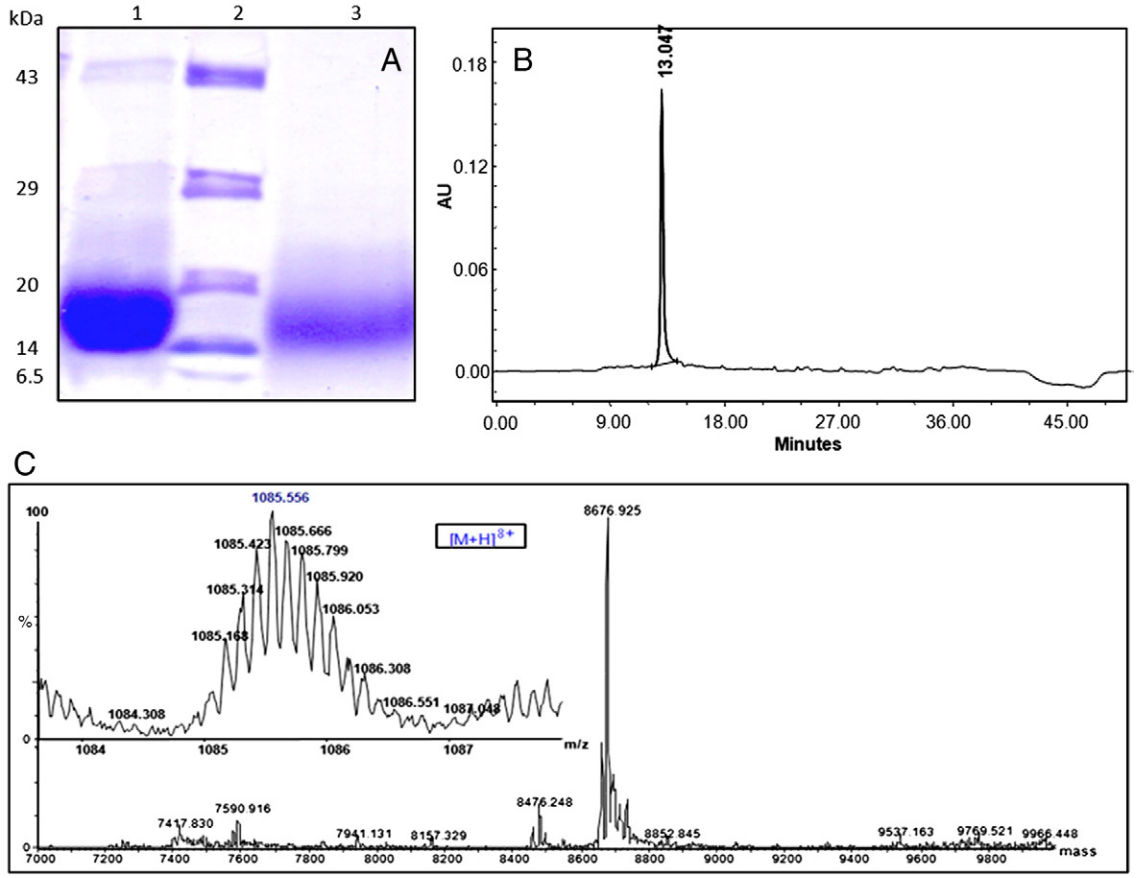


Fig. 3. (A) SDS–PAGE (15% T, 2.7% C) profile of purified HGIs. Lane 1, seed HGI-III; lane 2, molecular weight markers; and lane 3, purified rHGI. (B) RP–HPLC profile of purified rHGI and (C) ESI mass spectrum of rHGI. The inset shows an expansion of the molecular peak.

concentration of the disulphide bonds was calculated using a molar extinction coefficient of 13,600 M^{-1} cm^{-1} at 412 nm for NTB [47].

4. Results

4.1. Construction of pRSET-rHGI

The primers based on the amino acid sequence of HGI-III, the major BBI of horsegram, were used to amplify the coding sequences of genomic DNA. The only product obtained following amplification of genomic DNA with primers HGI-F/R annealed at 54–44 °C was a fragment of ~228 bp (Fig. 1A). Dideoxy DNA sequence analysis of the purified PCR product revealed that this fragment covered the entire coding sequence of HGI-III indicating the absence of introns. The deduced amino acid sequence was identical to that determined for HGI-III of horsegram [28] and differed by a single amino acid from the GenBank sequence (Accession No. AY049042). It is reported that genomic clones of BBI isolated from soybean (*Glycine max*) also do not contain any introns [48]. The PCR product encoding the 76 amino acids of HGI-III flanked by *Nde*I and *BamH*I was ligated into the *Pvu*II site of pRSETC using *T4* DNA ligase and confirmed by DNA sequencing and insert release by restriction digestion. A 44-amino acid stretch of the vector including the N-terminal $(His)_6$-tag before the coding sequence was removed by digestion with *Nde*I and allowed to self-ligate to obtain the expression plasmid designated pRSET-rHGI (Fig. 1B). The removal was confirmed by *BamH*I restriction digestion. The nucleotide sequence of the expression vector was determined, which showed the removal of the 44 amino acids. The DNA and translated amino acid sequence of HGI-III are shown in Fig. 1C.

4.2. Purification and biochemical characterization of rHGI

rHGI was expressed in *E. coli* BL21 (DE3) pLysS and solubilized by sonication. The total trypsin inhibitory activity of the crude cell lysate from a 1-L culture was $3.5 \pm 0.14 \times 10^5$ TIU with a specific activity of $1.54 \pm 0.44 \times 10^3$ TIU/mg protein. The inhibitor was purified using a single-step trypsin affinity chromatography. The elution profile indicated that rHGI eluted as a single peak when the pH was reduced to pH 3.0 (Fig. 2). The trypsin inhibitor fractions were pooled as shown in Fig. 2. The specific activity of the purified rHGI was $4.02 \pm 0.13 \times 10^3$ TIU/mg. Gelatin-embedded PAGE revealed that the rHGI inhibited bovine trypsin (Fig. 2, inset A). Native PAGE followed by incubation independently with either bovine trypsin or chymotrypsin showed that rHGI inhibited both the enzymes (Fig. 2, insets B and C). These results are consistent with that of the HGI-III purified from horsegram seeds, which is double-headed, inhibiting trypsin and chymotrypsin independently and simultaneously [26].

The estimated size of rHGI by SDS–PAGE was ~16,000 ± 1200 Da, which is consistent with its dimeric status in solution like the seed inhibitor (Fig. 3A). Size-exclusion HPLC on a BIOSEP-SEC-S 3000 column using 0.1 M Tris–HCl, pH 7.5, also revealed that rHGI eluted with a retention time of 9.293 min corresponding to a molecular mass of ~16,000 Da, which was in close agreement to similar to horsegram seed HGI-III (Figs. 4A and B). These results provide further evidence that rHGI like natural HGI-III in solution associates to form a dimer. This anomalous behavior of legume BBIs' is well documented. RP–HPLC of the purified rHGI indicated that the protein purified was homogenous (Fig. 3B). ESI–tandem MS indicated that the molecular mass of rHGI was 8676.925 Da, and the protein showed an isotopic pattern of $(M+H)^{8+}$ charge state (Fig. 3C). These results are in close agreement to that reported for HGI-III [26] and consistent with the theoretical value (8690 Da) deduced from the translated amino acid sequence. N-terminal sequence analysis by Edman degradation showed the sequence to be NH_2– MDHHQSTDEP…….consistent with that reported for horsegram HGI-III and that of the translated sequence.

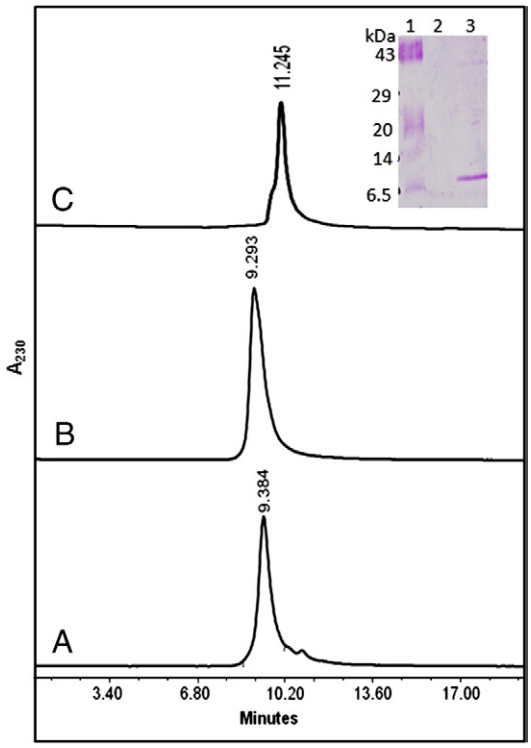


Fig. 4. Size-exclusion chromatography of seed HGI-III and rHGI. The purified proteins were dissolved in 0.1 M Tris–HCl, pH 7.5, and loaded on to a BIOSEP-SEC-S 3000 column preequilibrated in the same buffer and eluted at 1 mL/min. (A) HGI-III, (B) rHGI, and (C) K24A mutant of rHGI. Inset: SDS–PAGE (15% T, 2.7% C) followed by Western blotting and immune detection with anti-HGI-III. Lane 1, rainbow molecular weight markers; lane 2, bovine serum albumin; and lane 3, K24A mutant of HGI-III.

The chymotrypsin inhibitory activity of rHGI when assayed using the colorimetric substrate BTPNA was 100 ± 0.15 CIU/mg protein, which was much lower than the seed HGI-III (4572 CIU/mg protein). HGIs contain 14 half cystine residue linked by seven disulphide bonds. The number of disulphides in rHGI was estimated using the NTSB method. rHGI expressed in BL21 (DE3) pLysS showed the presence of seven disulphide bonds. These seven disulphide bridges are intramolecular as SDS–PAGE in the presence of β-mercaptoethanol shows a single polypeptide chain (Fig. 3A). These results suggest that the number of disulphide bonds was not responsible for the lowered chymotrypsin inhibitor activity. It is plausible that the low chymotrypsin inhibition is due to improper folding. rHGI was therefore overexpressed in the Origami strain of BL21 (DE3), a strain in which the mutations of both the thioredoxin reductase and glutathione reductase genes greatly enhance disulfide bond formation in the cytoplasm. The chymotrypsin inhibitory activity of the crude lysate increased to $2.9 \pm 0.13 \times 10^3$ CIU/mg protein. The yield of the purified protein was 1.0 ± 0.2 mg/L culture medium, when compared to 9.8 ± 2.3 mg/L for rHGI expressed in *E. coli* BL21 (DE3) pLysS. All further expressions were carried out using *E. coli* BL21 (DE3) pLys S.

4.3. pH and thermal stability of rHGI

Preincubation of rHGI for 60 min in the pH range of 3.0–9.0 had no effect on the trypsin inhibitor activity. Further 95% of the inhibitor activity was retained at all pH studied (Fig. 5A). Fig. 5B shows the thermal stability of rHGI. Heat treatment did not affect the trypsin inhibitory activity at 90 °C for 120 min. At 100 °C, rHGI showed a 20% decrease in the activity after 120 min (results not shown). These results are in agreement to that reported for other legume seed BBIs [1,5].

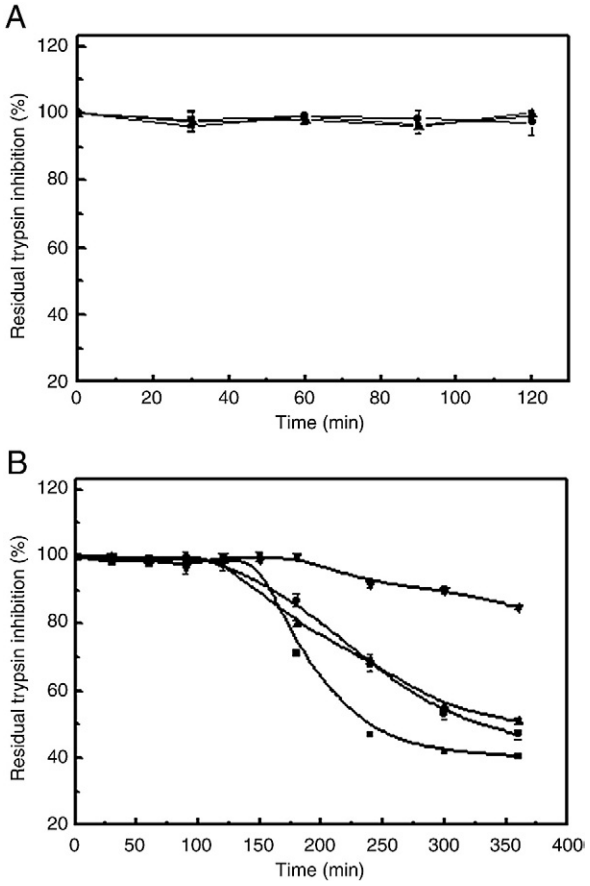


Fig. 5. (A) Effect of pH on the stability of rHGI. (■) pH 3.0, (●) pH 5.0, (▲) pH 7.0. (B) Thermal stability of rHGI and mutants at 90 ± 1 °C. (▼) rHGI, (■) K24A, (▲) pET-20b-HGI Δ76, and (●) pET-20b-HGI D75A. The inhibitors were incubated at 90 ± 1°C in a water bath. At regular time intervals, the residual trypsin/elastase inhibitor activity was determined.

4.4. Inhibitory properties of rHGI

The stoichiometry of inhibition against bovine pancreatic trypsin and chymotrypsin was assessed using BAPNA and BTPNA, respectively. Increasing concentrations of rHGI were incubated with a fixed concentration of the enzyme and the residual enzyme activity assayed. A linear extrapolation to obtain 100% inhibition indicated that rHGI bound to trypsin in a 1:0.9 molar ratio (Fig. 6), whereas there was no obvious stoichiometry with chymotrypsin from the titration pattern of the inhibitory activity (results not shown) similar to the HGIs isolated from horsegram seed [27].

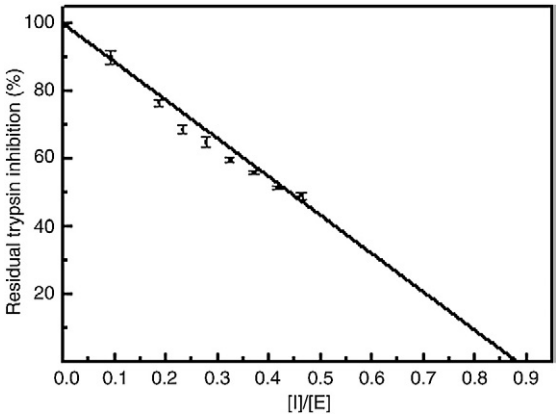


Fig. 6. Stoichiometric titration of bovine trypsin inhibition by rHGI. Increasing quantities of inhibitor were added to a fixed concentration of enzyme (2.5 nM). Residual enzyme activity was determined using BAPNA. Each point is the average of three assays.

The initial rates of reaction in the presence and absence of rHGI followed Michaelis–Menten kinetics (results not shown). The mode of rHGI inhibition was evaluated from the double reciprocal plots of trypsin/chymotrypsin titrated with different concentrations of their respective substrates. The results indicate that rHGI is a competitive inhibitor of both trypsin and chymotrypsin. The apparent K_i for trypsin inhibition from Dixon plots of the same data was $5 \pm 0.15 \times 10^{-8}$ and $6.1 \pm 0.13 \times 10^{-8}$ M for rHGI and refolded rHGI, respectively (Fig. 7A and Table 2). The K_i of HGI-III purified from

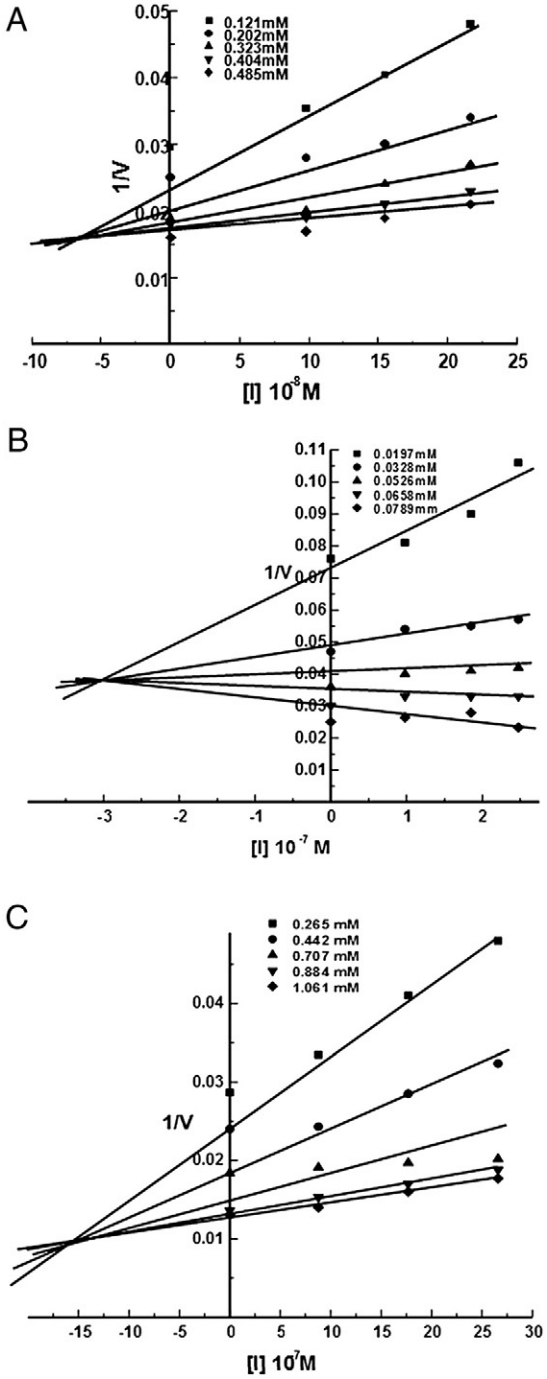


Fig. 7. Dixon plot for determining the dissociation constant (K_i) of rHGI for (A) bovine trypsin, (B) bovine chymotrypsin, and (C) K24A mutant against porcine pancreatic elastase. Data points are the average of three determinations.

Table 2
Dissociation constants (K_i) of HGI-III and rHGI for trypsin, chymotrypsin, and elastase inhibition.

Inhibitor	Bovine trypsin	Bovine chymotrypsin	Porcine pancreatic elastase
HGI-III	$8.7 \pm 0.13 \times 10^{-8}$ M	$3.9 \pm 0.16 \times 10^{-7}$ M	NA
rHGI	$6.1 \pm 0.13 \times 10^{-8}$ M ($5.0 \pm 0.15 \times 10^{-8}$ M)[a]	$3.0 \pm 0.15 \times 10^{-7}$ M ($3.4 \pm 0.12 \times 10^{-6}$ M)[a]	NA
K24A	NA	$2.6 \pm 0.13 \times 10^{-7}$ M	$1.58 \pm 0.14 \times 10^{-8}$ M ($4.0 \pm 0.13 \times 10^{-7}$ M)[a]
D75A	$4.8 \pm 0.11 \times 10^{-8}$ M	$3.4 \pm 0.18 \times 10^{-7}$ M	NA
Δ76	$3.8 \pm 0.23 \times 10^{-8}$ M	$2.2 \pm 0.20 \times 10^{-7}$ M	NA

NA, not applicable.
[a] Values in parenthesis are the values before refolding.

horsegram seed was $8.7 \pm 0.13 \times 10^{-8}$ M for trypsin, whereas the K_i toward chymotrypsin was $3.9 \pm 0.16 \times 10^{-7}$ M. The inhibitory constant of rHGI for chymotrypsin ($3.4 \pm 0.12 \times 10^{-6}$ M) was one order of magnitude below the activity of the natural inhibitor (Table 2). The consideration that this difference could result from some incorrectly folded protein coupled with the reported flexibility of the chymotrypsin domain prompted investigation on unfolding and refolding of rHGI. Refolding of rHGI led to a decrease in one order of magnitude for the chymotrypsin inhibitory constant (Fig. 7B and Table 2). In contrast, refolding had a very marginal effect on the K_i for trypsin inhibition (Table 2). These K_i values thus establish a very high affinity between these proteases and rHGI, in close agreement with the K_i values reported for other legume BBIs.

4.5. Site-directed mutagenesis of rHGI

The inhibitors of horsegram (HGIs) are single polypeptides of M_r ~8600 Da yet exists as a dimer in solution [27]. Kumar et al. [32] combining their observations on the monomeric status of the germinated seed HGGIs, chemical modification of HGI-III, and multiple sequence alignment of various legume BBIs concluded that a unique interaction involving the active site K^{24} and the C-terminal Asp was responsible for the dimerization of HGI-III. Therefore, to validate this hypothesis, site-directed mutagenesis of rHGI was carried out by Quik Change PCR method. Initially, the trypsin reactive site K^{24} was mutated to A. The mutant rHGI-K24A was expressed in *E. coli* BL21(DE3) pLys (yield of 6.2 mg/L culture broth) partially purified and refolded. The purified protein, however, lost its trypsin inhibitory activity but retained its chymotrypsin inhibitory potential (Table 2). This was not unexpected as the trypsin recognition site of rHGI is K^{24}. A is reported to be at the reactive site of BBIs that inhibit elastase. Therefore, rHGI-K24A was evaluated against porcine pancreatic elastase. The elastase inhibitory activity with its substrate was 4.8×10^3 EIU/mg protein. Kinetic analysis of the refolded protein indicated that it is a potent competitive inhibitor of elastase with a K_i of $1.58 \pm 0.14 \times 10^{-8}$ M, which was one order of magnitude more potent than the chymotrypsin inhibitory constant (Table 2 and Fig. 7C).

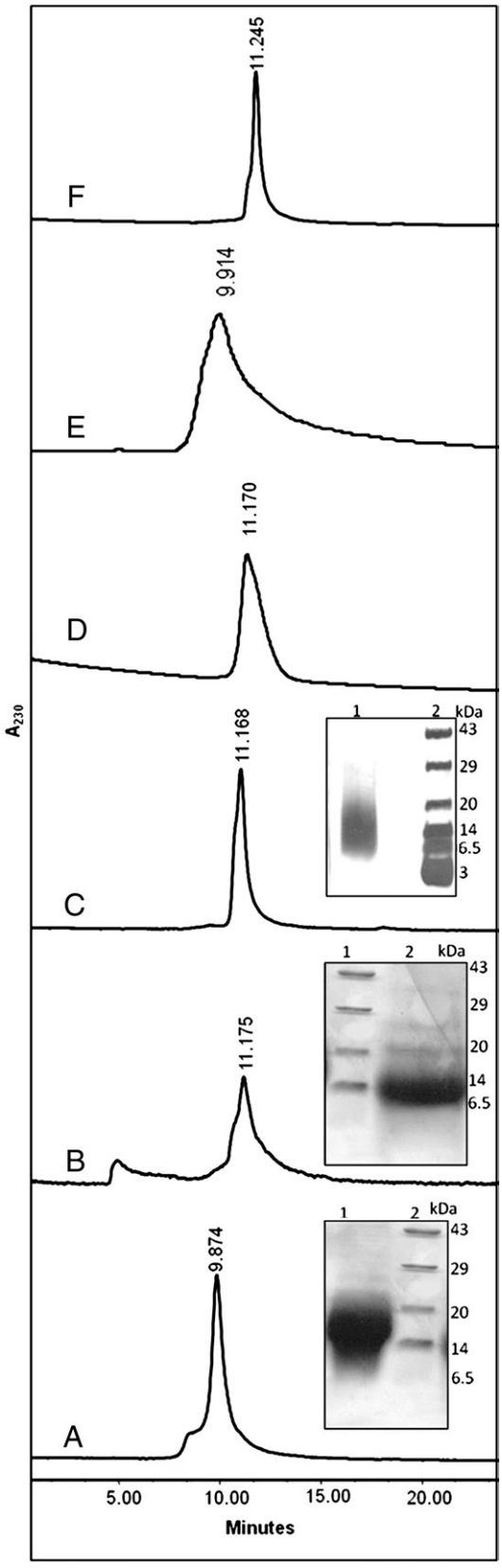
The self-association of rHGI-K24A mutant was evaluated by size-exclusion chromatography and SDS–PAGE, followed by Western blotting using anti-HGI-III antibodies available in our laboratory. The results shown in Fig. 4C indicate that the molecular mass of K24A mutant was ~8600 Da and, therefore, exists as a monomer in solution. These results further advocate that K^{24} is involved in the dimerization of the HGI-III. The purification of the K24A mutant was hindered by the fact it had lost its binding affinity to the trypsin–Sepharose matrix. Relatively little or no changes in the elastase inhibitory activity were observed up to 2 h for the K24A mutant. Although the stability of K24A mutant was comparable to rHGI up to 2 h, it lost 30% and 50% of its activity after 3 and 6 h of incubation at 90 °C, respectively (Fig. 5). These results indicate that the monomer was less stable than the dimer.

To abet in the purification of the expressed protein, rHGI was recloned using the vector pET20b to generate a $(His)_6$ fusion tag at the C-terminal end. The expression vector pET20b-HGI was used in all further experiments and mutation studies. pET20b-HGI-K24A was expressed in *E. coli* BL21(DE3) pLys and purified by Ni^{2+}-Sepharose chromatography. The yield was 6.2 mg/L culture broth. Size-exclusion chromatography and SDS–PAGE followed by protein staining showed that K24A $(His)_6$ tagged fusion protein moved as a protein of molecular mass ~16,000 ± 1000 Da (Fig. 8A). This was surprising and not expected.

The role of the C-terminal $D^{75\&76}$ was studied by carrying out a D75A and Δ76 mutation. The proteins were expressed in *E. coli* BL21 (DE3) pLys and purified by Ni^{2+}-Sepharose and size-exclusion chromatography. The yields were 6.8 ± 0.4 and 4 ± 0.9 mg/L culture broth, respectively. The purified proteins were refolded as described under the Methods section. The refolded D75A and Δ76 HGI mutant proteins inhibited both trypsin and chymotrypsin, and their inhibitory constants were similar to those of the refolded rHGI (Table 2). These mutants were, however, found to be less thermally stable when compare to rHGI (Fig. 5B). SDS–PAGE and size-exclusion studies indicate that both the D75A with the fusion tag and Δ76 mutants exist as monomers in solution (Figs. 8B and C). These results further advocate the role of $Asp^{75\&76}$ in the dimerization of HGI in solution. SDS–PAGE analysis of the purified $(His)_6$-tagged fusion protein probed with anti-HGI antibodies also suggests that D75A and Δ76 mutants are monomers in solution (results not shown).

rHGI K24A is a monomer (Fig. 4C); in contrast, pET-20b-HGI K24A exists as a dimer in solution (Fig. 8A). The major difference between rHGI K24A and pET-20b-HGI K24A is that the latter has a His_6 fusion tag. Either the His_6 through an ionic interaction with the negatively charged N-terminus of the second monomer or some other interaction is responsible for this dimerization. Size-exclusion studies at different pH were used to understand the role of His if any. pET-20b-HGI K24A eluted as a monomer at pH 5.0 (Fig. 8D). At pH 8.0 (results not shown) and 9.5 (Fig. 8E), the elution time corresponded to that of a dimer. At pH 5, His is protonated; therefore, if involved in dimerization through an ionic interaction, the protein should exist as a dimer, yet it is a monomer. At pH 8.0 and 9.5, wherein His is unprotonated, the protein retains its dimeric status (results not shown), suggesting that the His is probably not involved in dimerization. A previously reported molecular model of HGI-III dimer shows that the C-termini are located at the dimer interface, and the C-terminal Asp from one subunit forms a strong salt bridge with R^{31} of the other subunit, which is brought into this appropriate orientation by a hydrogen bond between K^{24} and D^{76} of opposite subunits [32]. Therefore, it is possible that in pET-20b-HGI K24A, this salt bridge (R^{31}–D^{76}) is not disrupted, resulting in the dimer. The role of C-termini Asp was studied by size-exclusion chromatography in the presence of $ZnSO_4$. pET-20b-HGI K24A elutes at 11.245 min in the presence of 1 mM $ZnSO_4$ (Fig. 8F). This increased retention time reckons a significantly reduced molecular weight. The conversion of pET-20b-HGI K24A to a monomer occurs through the indirect Zn^{2+} coordination with carboxylate side chain of nearby $Asp^{(D75/76)}$ mediated by the neighbouring His^{74}. Size-exclusion chromatography in the presence of $ZnSO_4$ has been used previously by us [32] to indirectly provide evidence that the Asp residues at the C-termini play a vital role in the dimerization of HGI-III. It is also possible that a substantial shift in the pK_a of $D^{75/76}$ due to the increase in length at the C-terminus by $(His)_6$ tag fusion reckoning the protonation at pH 5.0. Therefore, the interactions involving D75/76 are disrupted, and the protein is dissociated. The observed dissociation of pET-20b-HGI K24A to a monomer in the presence $ZnSO_4$ at pH 7.5 (Fig. 8F) in conjunction with the observation that pET-20b-HGI D75A with a similar His_6 fusion tag is a monomer (Fisg. 8B and C) points to the involvement of

C-terminal $D^{75/76}$ in the dimerization of this mutant. Site-directed mutations of HGI-III described clearly exemplify the pivotal role of the C-terminal $D^{75/76}$ in the dimerization previously demonstrated by chemical modifications [32].



5. Discussion

The results presented in this paper demonstrate the functional expression of the major BBI, HGI-III of horsegram seeds, and exemplify the pivotal role of the C-terminal Asp tail in the dimerization of HGI-III in solution. Previously, the identification of the role of individual amino acids in the structure and stability of the HGI-III dimer was accomplished using alternate approaches to site-directed mutagenesis like producing deletion variants of the inhibitor by germination, coupled with chemical modifications and homology modeling [32]. The self-association of HGI-III was attributed to a unique interaction that involved the ε-amino group of K^{24} (P1 of first reactive site) of one monomer and the carboxyl side chain of D^{76} of the other monomer. In this article, we have cloned and expressed HGI-III, and through a site-directed mutagenesis approach, we further exemplify that the self-association of HGI-III to form a dimer involves K^{24} and both the C-terminal D^{75} and D^{76}.

The expression and purification of HGI-III, the major BBI of horsegram, and its mutations have not been previously reported, although the isolation and purification of natural BBIs from horsegram seeds and its biochemical characteristics are well documented [27–32]. In this study, rHGI was overexpressed in a soluble form in *E. coli* BL21 cells. The target rHGI could be easily purified by a one-step trypsin affinity chromatography utilizing the strong binding potential to trypsin (Fig. 2). Unlike other BBIs that have been expressed as fusion proteins, rHGI was expressed with no fusion tags. The molecular mass of rHGI was 8676.25 Da, and the purity was >98% (Fig. 3). Further, the inhibitory activity assays suggested that rHGI like the natural HGI-III strongly inhibited both bovine trypsin and chymotrypsin (Figs. 7A and B). The dissociation constant of refolded rHGI for trypsin was $6.1 \pm 0.13 \times 10^{-8}$ M, which is similar to 8.25×10^{-8} M reported for natural HGI-III [27]. In contrast, the inhibition toward chymotrypsin was one order of magnitude below the activity of the natural inhibitor. This difference was due to incorrect folding as reckoned by the decreased K_i values after refolding (Table 2). The dissociation constant for chymotrypsin of the refolded rHGI was $3 \pm 0.15 \times 10^{-7}$ M similar to the seed HGI-III. A chemically synthesised soybean BBI gene cloned and expressed as a β-galactosidase fusion protein also showed that the dissociation constant of complexes with trypsin were similar to the natural BBI [35]. Prokaryotic expression of a rice BBI shows that the fusion protein has trypsin inhibitory activity but lacks chymotrypsin inhibitory activity [49]. The similarity of dissociation constants rHGI and seed HGI-III accompanied by competitive inhibition of trypsin and chymotrypsin affirms that the intramolecular disulphide bridges are in the correct orientation. In addition, the extreme thermal and pH stability exhibited by rHGI indicates the correct folding and disulphide formation. Further, the experimental measurements of seven disulphide bonds in rHGI are commensurate with the theoretical values. Data on the cloning and expression of legume BBI are sparse. The gene for buckwheat trypsin inhibitor was cloned and expressed in *E. coli* and shown to specifically inhibit the proliferation of IM-9 human B lymphoblastoid cells in a dose-dependent manner, but the dissociation constants were not determined [50]. Using chemical DNA synthesis, a fully active recombinant soybean BBI [35] was expressed in *E. coli*, and a cDNA coding for lentil trypsin/chymotrypsin BBI expressed in the methylotrophic yeast *Pichia pastoris* was functionally active [38].

Fig. 8. Size-exclusion chromatography of rHGI mutants expressed with $(His)_6$ fusion protein. The samples were dissolved in different buffers and loaded on to a BIOSEP-SEC-S 3000 column preequilibrated with respective buffers and eluted at 1 mL/min. (A) pET-20b-HGI K24A (0.1 M Tris–HCl pH 7.5), (B) pET-20b-HGI D75A (0.1 M Tris–HCl pH 7.5), (C) pET-20b-HGI Δ76 (0.1 M Tris–HCl pH 7.5), (D) pET-20b-HGI K24A (0.1 M Tris–HCl pH 5.0), (E), pET-20b-HGI K24A (0.1 M Tris–HCl pH 9.5), and (F) pET-20b-HGI K24A (1 mM $ZnSO_4$, in 0.1 M Tris-HCl pH 7.5). Insets in A, B, and C: SDS–PAGE (15% T, 2.7% C) profile of K24A, D75A, and Δ76 mutants, respectively.

The inhibitors of horsegram are thermally stable single polypeptides with molecular mass in the range ~8600 Da. SDS–PAGE and analytical gel filtration indicate the molecular mass to be ~16,000 Da [27], suggesting that they exist as dimers in solution. In a similar manner, rHGI with a molecular mass of 8676.92 Da shows a mass of ~16,000 Da by SDS–PAGE (Fig. 3A) and size-exclusion chromatography (Fig. 4A) and, therefore, like natural HGI-III, exists as a dimer in solution. Such self-association and anomalous behavior in SDS–PAGE resulting in a large overestimation of molecular mass has been reported for several legumes [51–54]. Similarly, an in vitro-synthesised BBI and related soybean inhibitor also exhibits the phenomenon of self-association [55].

We had earlier demonstrated that an electrostatic interaction involving the trypsin reactive site (K^{24}) and the C-terminal is the driving force for HGI-III self-association [32]. This self-association, together with other structural aspects, account for the stability of legume BBIs [19,30]. Previously, studies other than site-directed mutagenesis were used to deduce the crucial role of an electrostatic interaction between K^{24} and D^{76} in the self-association of HGI-III [32]. To conclusively prove this K24A mutation, rHGI was carried out and protein-purified. The lack of its binding affinity to trypsin Sepharose and the loss of trypsin inhibitor activity evidence the mutation. The dissociation constant of the complex with elastase (K_i, $1.58 \pm 0.14 \times 10^{-8}$ M) establishes a very high affinity. Size-exclusion chromatography, SDS–PAGE of the partially purified protein, and Western blotting with anti-HGI-III indicate a mass of ~8600 Da, clearly establishing the involvement of the reactive site K^{24} in the self-association of rHGI. The association of rHGI caused by this unique interaction between the two monomers is only of the monomer→dimer type with little or no higher forms present as is observed from size-exclusion studies (Fig. 4). A C-terminal $(His)_6$-tagged fusion protein of HGI (pETHGI) like rHGI is a dimer. Unexpectedly, the K24A mutant of this form was also a dimer. Size-exclusion studies either in the presence of $ZnSO_4$ or at pH 5.0 (Figs. 8D and F) led to the monomerization of K24A $(His)_6$ mutant. The disassociation of the pET-20b-HGI K24A into a monomer at pH 5.0 and in the presence of 1 mM $ZnSO_4$ together with the monomer status of pET-20b-HGI D75A and Δ76 mutant protein strengthens the premise that the C-termini $D^{75/76}$ play a more important role than K^{24} in dimer formation. The dimer model of seed HGI-III clearly discerns the role of the C-termini Asp residues in two contacts between the monomers: 1) a salt bridge (R^{31}–D^{76}) between the monomers and 2) a hydrogen bond with K^{24} to provide the required orientation for the formation of the salt bridge (R^{31}–D^{76}) [32]. The reported molecular model of HGI-III dimer indicates that the amino termini of the subunits are situated at the surface of the dimer, and any extension in this region would project into the solvent. This would not influence the stability of the dimer. In contrast, the carboxyl termini are located at the dimer interface and play an important role in the dimer stabilization. Therefore, any extension of C-termini would have consequences on the dissociation of the dimer as observed with the pET-20b-HGIK24A mutant. Xu et al. [56] report that the attachment of fusion sequences to coiled-coil proteins affects not only the thermal stability but also the oligomerization state. The monomer status of the pET-20b-HGI D75A and Δ76 mutant also points to the involvement of the C-terminal Asp residues in self-association. These results, in conjunction with our previously published data [32], unequivocally establish the role of an electrostatic interaction between K^{24} of one monomer and $D^{75/76}$ at the opposite face of the second monomer in self-association. NMR analysis of the self-association behavior of snail medic seeds BBI reveal that the residues involved are localized at opposite faces of the molecule, having the highest positive and negative potentials [17]. Most of the residues involved in self-association are highly conserved in BBIs from different seeds. Multiple sequence alignments of legume BBIs show that if the first reactive site is either K or R, they inhibit trypsin and tend to self-associate. In elastase inhibitors wherein K/R is replaced by A, they exist as monomers [32]. The K24A mutant of rHGI was converted to an elastase inhibitor (Fig. 7C) and, in solution, was a monomer (Fig. 4C), justifying the previous observation. The self-association of legume BBIs relate to their function in storage since such affinity is a requirement for the molecular packing in seeds [19]. The decreased thermal stability of the K24A mutant in comparison to rHGI further advocates that the dimeric form of the inhibitor is more stable than the monomer. The dimer HGI-III was also more thermostable than the monomeric HGGI-III isolated from germinated horsegram seeds [32].

The BBIs are promising models to study protein–protein interactions and to clearly distinguish between the structural and functional aspects using recombinant DNA techniques. Data on the cloning and expression of BBIs are limited. The cloning and heterologous expression of a functional rHGI provides a platform for production and systematic alteration of amino acid residues to explore the stability and mechanism of action and to unveil the fine specificity. The approach presented in this article has yielded on refolding in solution a fully active recombinant protein with seven disulphide bridges. In this study using a site-directed mutagenesis approach, we have demonstrated that the self-association of HGIs is indeed due to the electrostatic interaction between K^{24} of one monomer and $Asp^{75/76}$ of the second monomer, in agreement with our previous data [32]. Site-directed mutagenesis to understand the effects of the special pattern of seven disulphide bridges on the structure and stability of HGI-III is currently being investigated.

Acknowledgements

We are grateful to Dr. V. Prakash, Director, Central Food Technological Research Institute (CFTRI), Mysore, for his advice and valuable suggestions. The authors wish to thank Dr. A. G. Appu Rao, Head, Department of Protein Chemistry and Technology, CFTRI, Mysore, India, for his keen interest in this work. We also thank Dr. H. S. Savithri, Department of Biochemistry, Indian Institute of Science, for valuable suggestions. We thank Dr. H. Umesh Hebbar, Scientist, Department of Food Engineering, CFTRI, Mysore, for help with ultrafiltration. This investigation was supported by a grant from the Department of Science and Technology, New Delhi, India (SR/SO/BB-47/2002). D.G.M. acknowledges CSIR India for the award of a Senior Research Fellowship.

References

[1] M. Richardson, Seed storage proteins: the enzyme inhibitors, in: P.M. Dey (Ed.), *Methods in Plant Biochem*, 5, Academic Press, New York, NY, 1991, pp. 259–305.
[2] C.A. Ryan, Protease inhibitors in plants: genes for improving defences against insects and pathogens, Annu. Rev. Phytopathol. 28 (1990) 425–449.
[3] P.R. Shewry, J.A. Lucas, Plant proteins that confer resistance to pests and pathogens, Adv. Botan. Res. 26 (1997) 135–192.
[4] G. Chilosi, C. Caruso, C. Caporale, L. Leonardi, L. Bertini, A. Buzi, M. Nobile, P. Margo, V. Buonocore, Antifungal activity of a Bowman–Birk type trypsin inhibitor from wheat kernel, J. Phytopathol. 148 (2000) 477–481.
[5] M.J. Laskowski, I. Kato, Protein inhibitors of proteinases, Annu. Rev. Biochem. 49 (1980) 685–693.
[6] W. Bode, R. Huber, Natural protein proteinase inhibitors and their interaction with proteinases, Eur. J. Biochem. 204 (1992) 433–451.
[7] W. Bode, R. Huber, Structural basis of the endoproteinase–protein inhibitor interaction, Biochim. Biophys. Acta 1477 (2000) 241–252.
[8] V.V. Mosolov, T.A. Valueva, Proteinase inhibitors and their function in plants: a review, Prikl. Biokhim. Mikrobiol. 41 (2005) 261–282.
[9] T. Ikenaka, S. Norioka, Y. Tsunogae, A. Suzuki, T. Sone, K. Takahashi, I. Tanaka, T. Yamane, T. Ashida, S. Hara, Crystallization of Bowman–Birk type protease inhibitor (peanut) and its complex with trypsin, J. Biochem. 100 (1986) 243–246.
[10] Y. Birk, The Bowman–Birk inhibitor, Int. J. Pept. Prot. Res. 25 (1985) 113–131.
[11] P. Chen, J. Rose, R. Love, C.H. Wei, B.C. Wang, Reactive sites of an anticarcinogenic Bowman–Birk proteinase inhibitor are similar to other trypsin inhibitors, J. Biol. Chem. 267 (1992) 1990–1994.
[12] A. Suzuki, T. Yamane, T. Ashida, S. Norioka, S. Hara, T. Ikenaka, Crystallographic refinement of Bowman–Birk type protease inhibitor A-II from peanut (*Arachis hypogaea*) at 2.3 Å resolution, J. Mol. Biol. 234 (1993) 722–734.

[13] M.H. Werner, D.E. Wemmer, Three-dimensional structure of soybean trypsin/chymotrypsin Bowman–Birk inhibitor in solution, Biochemistry 31 (1992) 999–1010.
[14] R.H. Voss, U. Ermler, L.O. Essen, G. Wenzl, Y.M. Kim, P. Flecker, Crystal structure of the bifunctional soybean Bowman–Birk inhibitor at 0.28-nm resolution, structural peculiarities in a folded protein conformation, Eur. J. Biochem. 242 (1996) 122–131.
[15] J.R. Koepke, U. Ermler, E. Warkentin, G. Wenzl, P. Flecker, Crystal structure of cancer chemopreventive Bowman–Birk inhibitor in ternary complex with bovine trypsin at 2.3 Å resolution. Structural basis of Janus-faced serine protease inhibitor specificity, J. Mol. Biol. 298 (2000) 477–491.
[16] I.L.L. Sierra, L. Quillien, P. Flecker, J. Gueguen, S. Brunie, Dimeric crystal structure of a Bowman–Birk protease inhibitor from pea seeds, J. Mol. Biol. 285 (1999) 1195–1207.
[17] M. Catalano, L. Ragona, H. Molinari, A. Tava, L. Zetta, Anticarcinogenic Bowman Birk inhibitor isolated from snail medic seeds (*Medicago scutellata)*: solution structure and analysis of self-association behaviour, Biochemistry 42 (2003) 2836–2846.
[18] K.N. Rao, C.G. Suresh, Bowman–Birk protease inhibitor from the seeds of *Vigna unguiculata* forms a highly stable dimeric structure, Biochim. Biophys. Acta 1774 (2007) 1264–1273.
[19] J.A.R.G. Barbosa, L.P. Silva, R.C.L. Teles, G.F. Esteves, R.B. Azevedo, M.M. Ventura, S. M. de Freitas, Crystal structure of the Bowman–Birk inhibitor from *Vigna ungulata* seeds in complex with β-trypsin at 1.55 Å resolution and its structural properties in association with proteinases, Biophys. J. 92 (2007) 1638–1650.
[20] A.R. Kennedy, The Bowman–Birk inhibitor from soybean as an anticarcinogenic agent, Am. J. Clin. Nutr. 68 (1998) 1406–1412.
[21] Y.W. Chen, S.C. Huang, S.Y.L. Shiau, J.K. Lin, Bowman–Birk inhibitor abates proteosome function and suppresses the proliferation of MCF7 breast cancer cells through accumulation of MAP kinase phosphatase-1, Carcinogenesis 26 (2005) 1296–1306.
[22] A. Clemente, C. Domoney, Biological significance of polymorphism in legume protease inhibitors from the Bowman–Birk family, Curr. Protein. Pep. Sci. 7 (2006) 201–216.
[23] J.H. Ware, X.S. Wan, P. Newberne, A.R. Kennedy, Bowman–Birk inhibitor concentrate reduces colon inflammation in mice with dextran sulphate-induced ulcerative colitis, Digest. Dis. Sci. 44 (1999) 986–990.
[24] R.L. Gary, J.D. Julius, K. Seymour, D.L. James, A.R. Kennedy, J.H. Ware, Bowman–Birk inhibitor concentrate: a novel therapeutic agent for patients with active ulcerative colitis, Digest. Dis. Sci. 53 (2008) 175–180.
[25] B. Gran, N. Tabibzadeh, A. Martin, E.S. Ventura, J.H. Ware, G.X. Zhang, J.L. Parr, A.R. Kennedy, A.M. Rostami, The protease inhibitor, Bowman–Birk Inhibitor, suppresses experimental autoimmune encephalomyelitis: a potential oral therapy for multiple sclerosis, Mult. Scler. 12 (2006) 688–697.
[26] H.L. Blanca, C.C. Hsieh, B.O. de Lumen, Lunasin and Bowman–Birk protease inhibitor (BBI) in US commercial soy foods, Food Chem. 115 (2009) 574–580.
[27] Y.N. Sreerama, J.R. Das, D.R. Rao, L.R. Gowda, Double-headed trypsin /chymotrypsin inhibitors from horse gram (*Dolichos biflorus*): purification, molecular and kinetic properties, J. Food Biochem. 21 (1997) 461–477.
[28] B. Prakash, S. Selvaraj, M.R.N. Murthy, Y.N. Sreerama, D.R. Rao, L.R. Gowda, Analysis of the amino acid sequences of plant Bowman–Birk inhibitors, J. Mol. Evol. 42 (1996) 560–569.
[29] P.R. Ramasarma, A.G. Appu Rao, D.R. Rao, Role of disulphide linkages in structure and activity of proteinase inhibitor from horsegram (*Dolichos biflorus*), Biochim. Biophys. Acta 1248 (1995) 35–42.
[30] R.R. Singh, A.G.A. Rao, Reductive unfolding and oxidative refolding of a Bowman–Birk inhibitor from horsegram seeds *(Dolichos biflorus*): evidence for 'hyperreactive' disulfide bonds and rate-limiting nature of disulfide isomerization in folding, Biochim. Biophys. Acta 1597 (2002) 280–291.
[31] Y.N. Sreerama, L.R. Gowda, Antigenic determinants and reactive sites of a trypsin/chymotrypsin double headed inhibitor from horsegram (*Dolichos biflorus*), Biochim. Biophys. Acta 1343 (1997) 235–242.
[32] P. Kumar, A.G. Appu Rao, S. Hariharaputran, N. Chandra, L.R. Gowda, Molecular mechanism of dimerisation of Bowman–Birk inhibitors, pivotal role of Asp76 in the dimerization, J. Biol. Chem. 279 (2004) 30425–30432.
[33] L.P. Silva, R.B. Azevedo, P.C. Morais, M.M. Ventura, S.M. de Freitas, Oligomerization states of Bowman–Birk inhibitor by atomic force microscopy and computational approaches, Prot. Struct. Funct. Bioinformat. 61 (2005) 642–648.
[34] P. Kumar, Y.N. Sreerama, L.R. Gowda, Formation of Bowman–Birk inhibitors during the germination of horsegram (*Dolichos biflorus*), Phytochemistry 60 (2002) 581–588.
[35] P. Flecker, Chemical synthesis, molecular cloning and expression of gene coding for a Bowman–Birk type protease inhibitor, Eur. J. Biochem. 166 (1987) 151–156.
[36] N. Li, L.J. Qu, Y. Liu, Q. Li, H. Gu, Z.L. Chen, The refolding, purification, and activity analysis of a rice Bowman–Birk inhibitor expressed in *Escherichia coli*, Protein Expres. Purif. 15 (1999) 99–104.
[37] G. Vogtentanz, K.D. Collier, M. Bodo, J.H. Chang, A.G. Day, D.A. Estell, B.C. Falcon, G. Ganshaw, A.S. Jarnagin, J.T. Kellis, M.A. Kolkman, C.S. Lai, R. Meneses, J.V. Miller, H. de Nobel, S. Power, W. Weyler, D.L. Wong, B.F. Schmidt, A *Bacillus subtilis* fusion protein system to produce soybean Bowman–Birk protease inhibitor, Protein Expres. Purif. 55 (2007) 40–52.
[38] A. Clemente, D.A. Mackenzie, D.J. Jeenes, C. Domoney, The effect of variation within two domains on the activity of pea protease inhibitors from the Bowman–Birk class, Protein Expres. Purif. 36 (2004) 106–114.
[39] M.L. Kakade, N. Simons, I.E. Liener, An evolution of natural vs. synthetic substrate for measuring the antitryptic activity of soybean samples, Cereal Chem. 46 (1969) 518–526.
[40] M.M. Bradford, A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein–dye binding, Anal. Biochem. 72 (1976) 248–254.
[41] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[42] R. Felicioli, B. Garzelli, L. Vaccari, D. Melfi, E. Balestreri, Activity staining of protein inhibitors of protease on gelatin-containing polyacrylamide gel electrophoresis, Anal. Biochem. 244 (1997) 176–179.
[43] C.L. Fisher, G.K. Pei, Modification of a PCR-based site-directed mutagenesis method, Biotechniques 23 (1997) 570–574.
[44] H. Lineweaver, D. Burk, The determination of enzyme dissociation constants, J. Am. Chem. Soc. 56 (1934) 658–666.
[45] M. Dixon, The determination of enzyme inhibitor constants, Biochem. J. 55 (1953) 170–171.
[46] T.W. Thannhauser, Y. Konishi, H.A. Scheraga, Sensitive quantitative analysis of disulphide bonds in polypeptides and proteins, Anal. Biochem. 138 (1984) 181–188.
[47] G.L. Ellman, Tissue sulfhydryl groups, Arch. Biochem. Biophys. 82 (1959) 70–77.
[48] R.W. Hammond, D.E. Foard, B.A. Larkins, Molecular cloning and analysis of a gene coding for the Bowman–Birk protease inhibitor in soybean, J. Biol. Chem. 259 (1984) 9883–9890.
[49] L.J. Qu, J. Chen, M. Li, N. Pan, H. Okamoto, Z. Lin, C. Li, D. Li, J. Wang, G. Zhu, X. Zhao, X. Chen, H. Gu, Z. Chen, Molecular cloning and functional analysis of a novel type of Bowman–Birk inhibitor gene family in rice, Plant Physiol. 133 (2003) 560–570.
[50] Z. Zhang, Y. Li, C. Li, J. Yuan, Z. Wang, Expression of a buckwheat trypsin inhibitor gene in *Escherichia coli* and its effect on multiple myeloma IM-9 cell proliferation, Acta Biochim. Biophys. Sin. 39 (2007) 701–707.
[51] C. Wu, J.R. Whitaker, Purification and partial characterization of four trypsin/chymotrypsin inhibitors from red kidney beans (*Phaseolus vulgaris*, var. Linden), J. Agric. Food Chem. 38 (1990) 1523–1529.
[52] D. Bergeron, S.S. Nielsen, Partial characterization of trypsin inhibitors and N-terminal sequences of five trypsin inhibitors of great northern beans (*Phaseolous vulgaris*), J. Agric. Food Chem. 41 (1993) 1544–1549.
[53] S. Terada, S. Fujimura, S. Kino, E. Kimato, Purification and characterization of three proteinase inhibitors from *Canavalia lineata* seeds, Biosci. Biotechnol. Biochem. 58 (1994) 371–375.
[54] S.A. Godbole, T.G. Krishna, C.R. Bhatia, Changes in protease inhibitory activity from pigeon pea (*Cajanus cajan* (L) mill sp) during seed development and germination, J. Sci. Food Agric. 64 (1994) 87–89.
[55] D.E. Foard, P.A. Gutay, B. Ladin, R.N. Beachy, B.A. Larkins, In vitro synthesis of the Bowman–Birk and related soybean protease inhibitors, Plant Mol. Biol. 1 (1982) 227–243.
[56] C. Xu, L. Joss, C. Wang, M. Pechar, J. Kopecek, The influence of fusion sequences on the thermal stabilities of coiled-coil proteins, Macromol. Biosci. 2 (2002) 395–401.