CHAPTER 3

FUNCTIONAL EXPRESSION OF THE MAJOR BOWMAN-BIRK INHIBITOR OF HORSEGRAM-HGI-III
The Bowman-Birk inhibitor, is a prototype of a family of small proteinase inhibitors seen in legumes and many aspects of its chemistry and physiology, such as anticarcinogenic properties, highly differential specificity towards human granulocyte elastase and its role as a potent biological insecticide are of potential interest in medicine and agriculture. The symmetrical pattern of disulfide bridges, rigidifies inhibitory domains of BBI type proteinase inhibitors into a polycyclic, clearly arranged and highly conserved structural framework. 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 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 this chapter the results of cloning, expression and biochemical characterisation of a pure and fully active recombinant HGI-III (rHGI), the major BBI of horsegram seeds are presented.

RESULTS

Construction of pRSET-rHGI

Horsegram dry seed powder was prepared and defatted using five volumes of carbon tetrachloride and extraction overnight at 25ºC ± 2ºC. Genomic DNA was extracted from the defatted flour using genomic DNA extraction kit following manufacturer’s protocol (Qiagen, GmbH, Hilden, Germany). The extracted genomic DNA was analysed on 0.8 % agarose gel as described on Section 2.2.17 (Fig. 3.1).
Fig. 3.1 Agarose gel electrophoresis of horsegram genomic DNA. Lane 1: genomic DNA.

The yield and purity of the purified DNA was calculated. The $A_{260}/A_{280}$ ratio was 1.87. The yield was $1.65 \pm 0.31 \mu g$ of purified genomic DNA. Primers were designed based on the N-terminal and C-terminal sequence of HGI-III, the major BBI of horsegram and used to amplify the coding sequences of genomic DNA. The amplification was carried out as described in Section 2.2.21. The only product obtained following amplification of genomic DNA with primers HGI-F/R (Table 3.1) annealed at 54°- 44 °C was a fragment of ~228 bp (Fig 3.2).

Fig. 3.2 Agarose gel electrophoresis showing the 230-bp amplified product using the genomic DNA as template. Lane M: 100 bp gene ruler; Lane 1: Amplification product.
Fig. 3.3 (A) The nucleotide sequence of the amplified PCR product of HGI-III gene. (B) The translated sequence of the PCR product. The region in bold indicates the sequence of inhibitory loops.

Table 3.1 The list of oligonucleotides used for the amplification of HGI-III gene and construction of plasmid pRSET-rHGI

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGI-F</td>
<td>5`GATCATCATCAGTCAACTGATGAGCCCCTAGTCTAAACATGCCTGAG</td>
</tr>
<tr>
<td>HGI-R</td>
<td>5`ATCATCATGTGACACAGTTGTTTGGCAAGG3’</td>
</tr>
<tr>
<td>HGI-Nde-F</td>
<td>5`CTAGCTAGCCCATATGGGATCATCATCAGTCA3’</td>
</tr>
<tr>
<td>HGI-BamHI-R</td>
<td>5`CGCGGATCTTAAATCATCATGGAAG3’</td>
</tr>
</tbody>
</table>
Fig. 3.4 Agarose gel electrophoresis showing the 260-bp amplicon with the engineered restriction sites. Lane 1: 100 bp DNA ladder; Lane 2: Amplification product (~260 bp).

Fig. 3.5 Agarose gel electrophoresis profile of plasmid DNA isolated from putative clones. Lane 1: pRSET C vector control, Lanes 2-10: recombinant clones.

Dideoxy DNA sequence analysis of the purified PCR product (Fig. 3.2) revealed that this fragment covered the entire coding sequence of HGI-III indicating the absence of introns (Fig. 3.3 A). The deduced amino acid sequence (Fig. 3.3 B) agreed with that determined for HGI-III of horsegram (Prakash et al., 1996) and differed by a single amino acid from the GenBank sequence (Acc.No. AY049042). It has been reported that genomic clones of BBI isolated from soybean (Glycine max) also do not contain any introns.
The vector pRSETC was digested with \textit{Pvu}II restriction enzyme and digested product gel eluted. \textit{Nde}I and \textit{BamHI} sites were introduced to the 228 bp amplified product of HGI-III using primers HGI-\textit{Nde}-F and HGI-\textit{BamHI}-R. The 260 base pair PCR product (Fig. 3.4) encoding the 76 amino acids of HGI-III flanked by \textit{Nde}I and \textit{BamHI} was ligated into pRSETC digested with \textit{Pvu}II using \textit{T4} DNA ligase. The ligated product was transformed into chemically competent \textit{E. coli} DH5\(\alpha\) cells.

The ampicillin resistant transformants from LB-agar plates containing 100 \(\mu\text{g/mL}\) of ampicillin were picked up individually and screened for the presence of the 228 bp HGI gene. The plasmid DNA was isolated as described in Section 2.2.22 and analysed by agarose gel electrophoresis (Fig. 3.5). The plasmid DNA showing a mobility shift with respect to the control pRSET C vector DNA were picked up and screened for the release of HGI-III insert. The recombinants were confirmed by expression analysis, DNA sequencing and insert release by restriction digestion using \textit{BamHI} (Fig. 3.6 A). The resultant vector named pRSET-rHGI harbours an extra forty-four aminoacids at the N-terminal region. The forty-four amino acid stretch of the vector including the N-terminal (His)\(_6\)-tag prior to the coding sequence was removed by digestion with \textit{Nde}I. \textit{Nde}I site being introduced at the 5' end of HGI-III gene, restriction digestion would abet in the removal of the extra sequence. The digested product was allowed to self ligate and transformed into competent \textit{E. coli} DH5\(\alpha\) cells (Section 2.2.20) to generate the plasmid pRSET-rHGI (3025 bp) (Fig. 3.7). The transformants were identified by restriction digestion with \textit{BamHI}. The plasmids that had lost the extra amino acids linearise whereas the ones which carried the extra amino acids will have an insert release (Fig. 3.6 B). The nucleotide sequence of the expression vector pRSET-rHGI (3025 bp) was determined, which confirmed the removal of the forty-four amino acids.
Fig. 3.6 **Agarose gel electrophoresis of the plasmid DNA** digested with *BamHI*. (A) Restriction digestion of pRSET-rHGI (3157 bps) Lane M: 100 bp ladder; Lane 1: pRSET-rHGI (3157); Lane 2: pRSET-rHGI (3157) digested with *BamHI* showing an insert release. (B) Restriction digestion of pRSET-rHGI (3025 bps) Lane 1: 100 bp ladder; Lane 2: pRSET-rHGI (3025); Lane 3: pRSET-rHGI (3157). The arrow indicates the released product.

Fig. 3.7 **Schematic representation of the expression cassette in pRSETC vector to produce rHGI protein.**
Functional expression of rHGI

The expression vector pRSET-rHGI was transformed with *E. coli* BL21 (DE3) pLysS cells for the over-expression of rHGI (Fig. 3.8). 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. The supernatant and pellet were analysed for inhibitory activity. The trypsin inhibitory activity was vested in the supernatant of the cell lysate. The 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.

Purification of rHGI

The cell lysate was prepared in 0.05 M Tris-HCl (pH 8.2) containing 0.02 M CaCl$_2$, 0.1 M NaCl. Undissolved particulate matter was separated by centrifugation at 10,000 rpm for 15 min at 4 °C. The clear supernatant obtained was loaded on to a trypsin-sepharose column (12 × 3.4 cm), pre-equilibrated 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 washed at a flow rate 30 mL/h 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. Two mL fractions were collected and assayed for protein as well for the trypsin/chymotrypsin inhibitory activity. The elution profile indicated that rHGI eluted as a single peak when the pH was reduced to pH 3.0 (Fig.3.9). The active fractions were pooled as indicated, dialysed against water to remove the buffer ions and freeze-dried. The specific activity of the purified rHGI was $4.02\pm 0.13\times 10^3$ TIU/mg for anti-trypsin.
Fig. 3.8 SDS–PAGE (15% T, 2.7% C) profile of cell free extracts of pRSET-rHGI. Lane 1: rHGI induced with 0.3mM IPTG; Lane M: Molecular weight markers and Lane 2: uninduced rHGI. Arrow indicates expressed rHGI.

Fig. 3.9 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 30 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. The arrow shows the active inhibitor fractions that were pooled.
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). It is plausible that the low chymotrypsin inhibition is due to improper folding. rHGI was therefore over expressed 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±0.13×10^3 CIU/mg protein. The yield of the purified protein was 1.0±0.2 mg/L culture medium, when compared to 9.8±mg/L for rHGI expressed in *E. coli* BL21 (DE3) pLysS. All further expressions were carried out using *E. coli* BL21 (DE3) pLys S.

**Biochemical characterization of rHGI**

Homogeneity of the purified rHGI was examined by SDS-PAGE, Native-PAGE, RP-HPLC and N-terminal sequence analysis.

![Fig. 3.10](image)

**Fig. 3.10** *Native-PAGE (10% T, 2.7% C) profile of rHGI showing trypsin and chymotrypsin inhibitory activity.* (A) Gelatin-embedded native-PAGE showing trypsin inhibition. The gels were stained for (B) trypsin inhibitory activity and (C) chymotrypsin inhibitory activity with APNE and (D) protein
The purified inhibitor was electrophoresed on native PAGE (Section 2.2.6), pH 8.3 (10 % T, 2.7 % C) and detected by staining with CBB R-250. Native PAGE followed by incubation independently with either bovine trypsin or chymotrypsin showed that rHGI inhibited both the enzymes (Fig. 3.10 A, B & C). The recombinant inhibitor was homogenous and migrated as single species both by protein staining (Fig. 3.10 D) and by specific staining for trypsin inhibitory activity (Fig. 3.10 B) and chymotrypsin inhibitory activity (Fig. 3.10 C). The purified protein was dissolved in 0.1 % TFA and applied on an RP-HPLC column (Waters Symmetry Shield™ RP18 150 mm × 4.6 mm (i.d), 5 µ column) and eluted using a binary gradient of 0.1 % TFA and 70 % acetonitrile in water containing 0.05 % TFA. RP-HPLC profile of the purified rHGI indicated that the protein was homogenous (Fig. 3.11).

![RP-HPLC profile of purified rHGI using a binary gradient of 0.1 % TFA and 70 % acetonitrile in water containing 0.05 % TFA.](image)

**Fig. 3.11** RP-HPLC profile of purified rHGI using a binary gradient of 0.1 % TFA and 70 % acetonitrile in water containing 0.05 % TFA.
Fig. 3.12 RP-HPLC elution profile of the PTH- amino acids released from rHGI in the first six cycles of Edman degradation by automated gas phase sequencing. The eluted peaks were compared with the standard profile to deduce the sequence. (A) Standard PTH aminoacids; B-D cycles 1-3.
RP-HPLC elution profile of the PTH- amino acids released from rHGI in the first six cycles of Edman degradation by automated gas phase sequencing. The eluted peaks were compared with the standard profile to deduce the sequence. E-G cycles 4-6.
N-terminal sequence analysis

The purified rHGI was subjected to N-terminal sequence analysis on an automated gas phase sequenator. The release of a single N-terminal amino acid residue Met (M) for rHGI evidenced the purity of the inhibitor. N-terminal sequence analysis by Edman degradation showed the sequence to be NH$_2$- MDHHQSTDEP.......(Fig. 3.12) consistent with that reported for horsegram HGI-III and that of the translated sequence (Fig. 3.3 B).

Molecular weight determination

The apparent size of rHGI was evaluated by SDS-PAGE and size exclusion chromatography. The size of rHGI by SDS-PAGE was ~16 ±1.2 kDa (Fig. 3.13 Lane 3), which is consistent with its dimeric status in solution like the seed inhibitor (Fig. 3.13 Lane 1). Size-exclusion-HPLC on a BIOSEP-SEC-S 3000 column using 0.1 M Tris-HCl, pH 7.25 also revealed that rHGI eluted with a retention time of 9.293 min corresponding to a molecular mass of ~16 kDa, which was in close agreement to horsegram seed HGI-III (Fig 3.14). These results provide further evidence that rHGI like horsegram seed HGI-III in solution associates to form a dimer. This anomalous behaviour of legume BBIs is well documented. ESI-tandem MS indicated that the molecular mass of rHGI was 8676.925 Da (Fig. 3.15) and the protein showed an isotopic pattern of (M+H)$^{8+}$ charge state (Fig. 3.16). These results are in close agreement to that reported for seed HGI-III and consistent with the theoretical value (8.690 kDa) deduced from the translated amino acid sequence (Fig. 3.3 B).
Fig. 3.13 *SDS-PAGE (15%T, 2.7%C) profile of purified HGI*s. Lane 1: Seed HGI-III, Lane 2: Molecular weight markers and Lane 3: purified rHGI.

Fig. 3.14 *Size-exclusion chromatography of seed HGI-III and rHGI*. The purified proteins were dissolved in 0.1M Tris-HCl, pH 7.5 and loaded on to a BIOSEP-SEC-S 3000 column pre-equilibrated in the same buffer and eluted at 1 mL/min. (A) HGI-III; (B) rHGI.
Fig. 3.15 ESI mass spectrum of rHGI showing the molecular mass of 8676.925 Da.

Fig. 3.16 An expansion of the molecular peak of rHGI of Fig. 3.15 showing an isotopic pattern of (M+H)^8+ charge state.
Estimation of disulphide bonds

Disulphide bridges are well conserved in BBIs. Therefore to estimate the number of disulphide bridges in rHGI, these bridges were initially cleaved quantitatively to free thiol using excess sodium sulphite, pH 9.5 at 37 °C. To denature the rHGI, guanidine hydrochloride (6 M) was added to the inhibitor solution, which makes the disulphide bonds accessible. The determination of free thiol groups released as a result of the denaturation of disulphide bonds was evaluated using NTSB (as described in section 2.2.39) in presence of excess sodium sulphite. The number of disulphide bonds reacted per molecule of the inhibitor, followed by the change in absorbance at 412 nm was calculated to be seven (Fig. 3.17). This corresponds to the presence of fourteen half cystine residues per molecule of rHGI. The number of disulphide bonds of the reference protein Ribonuclease A was also determined to test the validity of the analytical procedure, and found to be four which agrees well with the reported values (Fig. 3.17).

![Assay for disulphide bonds of rHGI](image)

**Fig. 3.17 Assay for disulphide bonds of rHGI.** The assay was carried out by following the change in absorbance at 412 nm due to NTB produced in the reaction. (▲) Ribonuclease A, (●) rHGI, (■) HGI-III.
pH and thermal stability of rHGI

BBI proteins are stable to extreme conditions like heat and can withstand wide range of pH. To understand whether rHGI also retains these properties, pH and thermal stability studies were conducted. Preincubation of rHGI for 60 mins 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. 3.18). The thermal stability studies of rHGI showed that heat treatment did not affect the trypsin inhibitory activity at 90 °C for 120 mins. At 100 °C, rHGI showed a 20% decrease in the activity after 120 min (Fig. 3.19). These results are in agreement to that reported for other legume seed BBIs (Richardson, 1991, Laskowski et al., 1980).

Fig. 3.18 Effect of pH on the stability of rHGI. (−■−) pH 3.0; (−•−) pH 5.0; (−▲−) pH 7.0. The residual trypsin inhibitor activity was determined.
Fig. 3.19 Thermal stability of rHGI. (■) 90 °C and (•) 100 °C. rHGI was incubated at 90±1 °C and 100±1 °C. At regular time intervals the residual trypsin inhibitor activity was determined.

**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. 3.20 A), whereas there was no obvious stoichiometry with chymotrypsin (Fig. 3.20 B) from the titration pattern of the inhibitory activity similar to the HGIs isolated from horsegram seed [Sreerama et al., 1997].
Fig. 3.20 (A) **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.

Fig. 3.20 (B) **Stoichiometric titration of bovine chymotrypsin inhibition by rHGI.** Increasing quantities of inhibitor were added to a fixed concentration of enzyme (5 nM). Residual enzyme activity was determined using BTPNA. Each point is the average of three assays.
The initial rates of reaction in the presence and absence of rHGI followed Michaelis Menten kinetics (Fig. 3.21). The mode of rHGI inhibition was evaluated from the double reciprocal plots (Fig. 3.22) of trypsin/chymotrypsin titrated with different concentration of their respective substrates. The results indicate that rHGI is a competitive inhibitor of both trypsin and chymotrypsin (Fig. 3.22). The apparent $K_i$ for trypsin inhibition from Dixon plots of the same data was $5\pm0.15 \times 10^{-8}$ M for rHGI (Fig. 3.23 & Table 3.2). The $K_i$ of HGI-III purified from horsegram seed was $8.7\pm0.13 \times 10^{-8}$ M for trypsin whereas the $K_i$ towards chymotrypsin was $3.9\pm0.16 \times 10^{-7}$ M. The inhibitory constant of rHGI for chymotrypsin ($3.4\pm0.12 \times 10^{-6}$ M) (Fig. 3.23 B) was one order of magnitude below the activity of the natural inhibitor (Table 3.2). 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.

### Table 3.2: Dissociation constants ($K_i$) of HGI-III and rHGI for trypsin, and chymotrypsin inhibition

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Bovine trypsin</th>
<th>Bovine chymotrypsin</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGI-III</td>
<td>$8.7\pm0.13 \times 10^{-8}$ M</td>
<td>$3.9\pm0.16 \times 10^{-7}$ M</td>
</tr>
<tr>
<td>rHGI</td>
<td>$5.0\pm0.15 \times 10^{-8}$ M</td>
<td>$3.4\pm0.12 \times 10^{-5}$ M</td>
</tr>
<tr>
<td>rHGI refolded</td>
<td>$6.1\pm0.13 \times 10^{-8}$ M</td>
<td>$3.0\pm0.15 \times 10^{-7}$ M</td>
</tr>
</tbody>
</table>
Fig. 3.21 Michaelis-Menten plot showing the effect of substrate concentration on the activity of (A) bovine trypsin (B) bovine chymotrypsin in the presence of various concentrations of rHGI. Data points are average of three determinations.
Fig. 3.22 Lineweaver-Burk plot of data from Fig. 3.21 showing competitive inhibition by rHGI for (A) trypsin (B) chymotrypsin. Data points are average of three determinations.
Fig. 3.23 Dixon plot for determining the dissociation constant ($K_i$) of rHGI. (A) bovine trypsin. (B) bovine chymotrypsin. Data points are the average of three determinations.
Refolding of the purified rHGI in solution was carried out as reported for CNBr cleaved recombinant soybean BBI (Flecker, 1987). Refolding of rHGI led to an increase in one order of magnitude for the chymotrypsin inhibitory constant (Table 3.2) in close agreement to the natural inhibitor. In contrast refolding had very marginal effect on the $K_i$ for trypsin inhibition (Table 3.2). These $K_i$ values thus establish a very high affinity between these proteases and rHGI, close in agreement with the $K_i$s reported for other legume BBIs.

**Simultaneous and independent inhibition of trypsin/chymotrypsin by rHGI**

The simultaneous inhibition of trypsin and chymotrypsin by rHGI were evaluated. Competitive inhibition studies were performed with the purified rHGI. In the first experiment, the chymotrypsin inhibitory activity was measured in the presence of varying amounts of trypsin. Initially rHGI was incubated with various concentrations of trypsin (0.5 μM to 3 μM) for 10 min at 37 °C and inhibition of chymotrypsin measured using BTPNA as substrate. Similarly, after incubating the inhibitor with various concentrations of chymotrypsin (0.5 μM to 3 μM) for 10 min at 37 °C, the effect on the inhibition of trypsin was measured using BAPNA as substrate. The competitive studies revealed that rHGI inhibits chymotrypsin even in presence of trypsin (Fig. 3.24) and viceversa (Fig. 3.25). The competitive binding studies point towards the independent and simultaneous binding of rHGI to the cognate enzymes. 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 (Sreerama et al., 1997).
Fig. 3.24 **Chymotrypsin inhibition by rHGI in the presence of varying concentrations of trypsin.** Data points are average of three determinations.

Fig. 3.25 **Inhibition of trypsin by rHGI in the presence of varying concentrations of chymotrypsin.** Data points are the average of three determinations.
Cross reactivity of rHGI with anti HGI-III

Cross-reactivity of the rHGI was studied using antibodies raised against HGI-III, the major inhibitor present in the dormant horsegram seeds (Sreerama and Gowda, 1997). The purified rHGI was immobilized on nitrocellulose membrane and subjected to immuno-detection. The blot was developed with a solution of BCIP and NBT to ascertain the cross reactivity. Positive antigen-antibody reaction was detected with rHGI (Fig. 3.26). HGI-III served as the positive control and BSA served the negative control, which showed no cross reactivity (Fig. 3.26).

![Figure 3.26](image.png)

Fig. 3.26 Dot-Blot analysis of recombinant horsegram inhibitor using antibodies raised against HGI-III. (1) rHGI (2) BSA and (3) HGI-III.

Discussion

Protease inhibitors are found in various plants and have been studied in Leguminosae, Gramineae and Solanaceae. The Bowman-Birk family and Kunitz family of proteinase inhibitors discovered first in soybean seeds have been studied extensively. BBIs are classified as double-headed inhibitors because the protein contains two reactive site domains within the same polypeptide. BBI’s are recognised to have a trypsin (Lys-Ser) and a chymotrypsin (Phe/Leu-Ser) inhibitory site (Birk, 1961), or two trypsin (Arg-
Ser) inhibitory sites (Hwang et al., 1977). The occurrence of multiple low molecular weight protease inhibitors among members of the legume family (Chu and Chi, 1965; Wilson and Laskowski, 1973) as well as distantly related plants (Bryant et al., 1976) suggests that the genes may have been conserved throughout the evolution of several plant species. Isolation of the genes encoding these protease inhibitors have provided information on their number and organization as well as probes for studying the regulation of their expression. Hammond et al., (1984) constructed and characterized a cDNA clone corresponding to the BBI. DNA sequence analysis confirmed that this sequence corresponds to the BBI mRNA. Although there is extensive sequence homology between the BBI cDNA and genomic clones, they differ slightly in their deduced amino acid sequences (Hammond et al., 1984). The gene size and coding regions of BBIs are small and devoid of introns (Boulter, 1993). The availability of these inhibitors may allow engineering of transgenic crops inherently resistant to these insect pests. These aspects therefore encourage the identification and characterization of novel protease inhibitor genes from diverse plant species. Besides their natural biological functions, protease inhibitors might also be useful in treating human pathologies such as inflammation, haemorrhage (Oliva et al., 2000) and cancer (Kennedy, 1998).

Recently, with the rapid development of biotechnology many studies on gene cloning and the expression of protease inhibitors have been reported (Li et al., 1999, Chen et al., 1997). A putative rice trypsin/chymotrypsin inhibitor of the Bowman–Birk family, about 20 kDa, was expressed in *E. coli* as a fusion protein bearing an N-terminal (His)_6_ purification tag (Li et al., 1999). To reveal the inhibitory activities towards proteases and tumour cells, the open-reading frame of the buckwheat trypsin inhibitor gene was chemically synthesized and expressed in *E. coli*
M15 (Zhang et al., 2007). The cDNA cloning of BBI from *Apios americana* tubers also has been reported (Zhang et al., 2008). The reports on the cloning and expression of soybean BBI also demonstrates that in almost all the strategies used, the clones were expressed as fusion proteins and required an activation process for the functional expression. The results presented here demonstrate the functional expression of the major BBI, HGI-III of horsegram seeds and the biochemical characterisation of the recombinant inhibitor. The expression and purification of HGI-III has not been previously reported although the isolation and purification of four BBIs from horsegram seeds and its biochemical characteristics have been documented (Sreerama et al., 1997, Prakash et al., 1996, Sreerama and Gowda, 1997, Kumar et al., 2004). In this study, the amplified DNA fragment produced by PCR reaction using genomic DNA isolated from horsegram seed as template was of expected size (~228 bp) covering the entire HGI-III gene. The DNA sequence analysis of the plasmid (pRSET-rHGI) indicated a coding region of 228 bp devoid of introns. None of the BBI genes sequenced so far has any introns; the isolated sequenced cowpea trypsin inhibitor gene is also devoid of introns (Lawrence et al., 2001). The amino acid sequence deduced for rHGI was in agreement to the amino acid sequence determined earlier (Prakash et al., 1996). rHGI was over expressed in a soluble form in *E. coli* BL21 cells (Fig. 3.8). The target rHGI could be easily purified by a single step trypsin affinity chromatography utilising the strong binding potential to trypsin (Fig. 3.9). Unlike other BBIs that have been expressed as fusion proteins, rHGI was expressed devoid of any fusion tags. The strategy used here in creating rHGI without the extra sequence (His tag fusion) could be used for other small proteins to avoid the interference of the fusion tags in the function, stability and other molecular interactions of the expressed proteins. Fusion tags from the vector have
been reported to interact and interfere with the thermal stability and oligomerisation status in the case of coiled coil proteins (Xu et al., 2002). The purification strategy used in our study yields ~98 % pure recombinant protein (Fig. 3.9) in a single step purification. The requirement of other proteases to cleave the fusion tag proteins, a strategy used in most of the recombinant proteins was avoided in the present study used. This was made possible by utilising the strong binding property between rHGI and trypsin for purification.

The molecular mass of rHGI was 8.676 kDa (Fig. 3.15) and the purity was >98 % (Fig. 3.11). Further the inhibitory activity assays suggested that rHGI like the natural seed HGI-III strongly inhibited both bovine trypsin and chymotrypsin (Fig. 3.10). The BBIs of horsegram have two separate reactive sites for inhibition of trypsin and chymotrypsin and are independent of each other. Soybean C-II inhibits both trypsin and chymotrypsin at a single reactive site (Odani and Ikenaka, 1977). The BBI, A-II isolated from peanut inhibits trypsin at two reactive sites, of which one site can also inhibit chymotrypsin (Norioka et al., 1982). The stoichiometry of inhibition against bovine pancreatic trypsin and chymotrypsin was assessed using BAPNA and BTPNA respectively. Inhibition of trypsin at fixed concentrations was studied by varying the rHGI concentrations to deduce the stoichiometric relation. Although a near 1:1 stoichiometry was observed for trypsin no such stoichiometry was obtained for chymotrypsin. No obvious stoichiometry for chymotrypsin binding was also observed in BBIs isolated from other legumes (Terada et al., 1994a; Norioka et al., 1982; Kimura et al., 1994; Godbole et al., 1994).

The competitive binding studies of rHGI with bovine trypsin and chymotrypsin revealed the simultaneous and independent inhibition of rHGI
towards these proteases similar to that reported for HGI-III (Fig. 3.24 and Fig. 3.25). The Adzuki bean inhibitor II inhibits trypsin and chymotrypsin at different sites, but its complex with chymotrypsin fails to give any trypsin inhibitory activity (Yoshikawa et al., 1979). Five BBIs from peanut (*Arachis hypogaea*) (Norioka et al., 1982) inhibit bovine trypsin and chymotrypsin at molar ratios of 1:2 and 1:1 respectively, yet the inhibition of trypsin and chymotrypsin are not independent of each other. Tur-Sinai et al., (1972) reported that the complex of their peanut inhibitor with trypsin lacked chymotrypsin-inhibiting activity and the complex with chymotrypsin no longer had trypsin inhibiting activity. The dissociation constant of rHGI and refolded rHGI for trypsin were similar to that reported for natural HGI-III (Sreerama et al., 1997). In contrast the inhibition towards 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 3.2). The dissociation constant for chymotrypsin of the refolded rHGI was $3\pm0.15\times10^{-7}$ M in close agreement to the natural HGI-III (Fig. 3.23). 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 are similar to the natural BBI (Flecker, 1987). Prokaryotic expression of a rice BBI shows that the fusion protein has trypsin inhibitory activity but lacks chymotrypsin inhibitory activity (Qu et al., 2003).

The similarity of dissociation constants of rHGI and seed HGI-III accompanied by competitive inhibition of trypsin and chymotrypsin (Fig.3.22) affirms that the intramolecular disulphide bridges are in the correct orientation. In addition the extreme thermal (Fig. 3.19) and pH stability (Fig. 3.18) exhibited by rHGI indicates the correct disulphide formation. Further the experimental measurements of seven disulphide
bonds in rHGI are commensurate with the theoretical values (Fig 3.17). BBIs are extra ordinary proteins which are able to survive the cooking and digestive process and reach the colon in the active stage. The thermal and pH stability study shows that the expressed BBI also retained this capability. Considering the positive therapeutic benefits of BBIs and their relative stability at cooking temperature (Birk, 1985) and the acidic digestive system in humans (Yavelow et al., 1983), BBI levels have been determined in commercial foods (Blanca et al., 2009). Soybean BBIs are reported to survive faecal fermentation and are active anticarcinogens (Marin-Manzano et al., 2009). Polyclonal antibodies raised against the major inhibitor, HGI-III cross reacts with the rHGI (Fig. 3.26) suggesting that the antigenic determinants are similar in both. 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 β lymphoblastoid cells in a dose dependent manner, but the dissociation constants were not determined (Zhang et al., 2007). Using chemical DNA synthesis, a fully active recombinant soybean BBI (Flecker, 1987) was expressed in *E. coli* and a cDNA coding for lentil trypsin/chymotrypsin BBI expressed in the methylotrophic yeast *Pichia pastoris* was functionally active (Caccialupi et al., 2008).

The rHGI inhibitor similar to that of dry seed inhibitor exhibited molecular masses of ~16 kDa by SDS-PAGE and analytical gel filtration, although ESI-MS analysis showed them to have a mass of ~ 8 kDa. This large overestimation of molecular weights has been attributed to the legume BBIs existing in a state of equilibrium between monomer-dimer-trimer forms, the dimer being the predominant form. (Wu and Whitaker, 1990; Terada et al., 1994a; Godbole et al., 1994). However such an overestimation of the molecular mass was not observed in germinated HGIs suggesting that they
exist only as monomers (Kumar et al., 2002). Similarly an in vitro synthesised BBI and related soybean inhibitor also exhibits the phenomenon of self association (Foard et al., 1982). In onion trypsin inhibitor, the molecular mass is 7.6426 kDa with respect to MALDI-TOF analysis but SDS-PAGE analysis shows a molecular mass of 18 kDa and 8 kDa in the absence of reducing agent whereas the inhibitor may exist as a mixture of monomer and dimer (Deshimaru et al., 2003).

The BBIs are promising models for studies on protein-protein interactions and to clearly distinguish between structural and functional aspects using recombinant DNA techniques. Data on the cloning and expression of BBIs is limited. The cloning and heterologous expression of a functional rHGI provides a platform for production and systematic alteration of amino acid residues to expedite the stability, mechanism of action and unveil the fine specificity. The approach presented here has yielded a pure and fully active recombinant protein with seven disulphide bridges. The functional expression of rHGI has paved the way to study mechanisms at the molecular level.