# **Molecular Mechanism of Dimerization of Bowman-Birk Inhibitors**

PIVOTAL ROLE OF  $\mathrm{ASP}^{76}$  IN THE DIMERZATION\*

Received for publication, March 17, 2004, and in revised form, April 27, 2004 Published, JBC Papers in Press, April 28, 2004, DOI 10.1074/jbc.M402972200

# Pradeep Kumar‡§, A. G. Appu Rao‡, Sridhar Hariharaputran¶, Nagasuma Chandra¶, and Lalitha R. Gowda‡ $\parallel$

From the ‡Department of Protein Chemistry and Technology, Central Food Technological Research Institute, Mysore 570 020 and the ¶Bioinformatics Centre, Indian Institute of Science, Bangalore 560 012, India

Horsegram (Dolichos biflorus), a protein-rich leguminous pulse, is a crop native to Southeast Asia and tropical Africa. The seeds contain multiple forms of Bowman-Birk type inhibitors. The major inhibitor HGI-III, from the native seed with 76 amino acid residues exists as a dimer. The amino acid sequence of three isoforms of Bowman-Birk inhibitor from germinated horsegram, designated as HGGI-I, HGGI-II, and HGGI-III, have been obtained by sequential Edman analyses of the pyridylethylated inhibitors and peptides derived therefrom by enzymatic and chemical cleavage. The HGGIs are monomers, comprising of 66, 65, and 60 amino acid residues, respectively. HGGI-III from the germinated seed differs from the native seed inhibitor in the physiological deletion of a dodecapeptide at the amino terminus and a tetrapeptide, -SHDD, at the carboxyl terminus. The study of the state of association of HGI-III, by size-exclusion chromatography and SDS-PAGE in the presence of 1 mM ZnCl<sub>2</sub>, has revealed the role of charged interactions in the monomer  $\leftrightarrow$  dimer equilibria. Chemical modification studies of Lys and Arg have confirmed the role of charge interactions in the above equilibria. These results support the premise that a unique interaction, which stabilizes the dimer, is the cause of selfassociation in the inhibitors. This interaction in HGI-III involves the  $\epsilon$ -amino group of the Lys<sup>24</sup> (P<sub>1</sub> residue) at the first reactive site of one monomer and the carboxyl of an Asp<sup>76</sup> at the carboxyl terminus of the second monomer. Identification of the role of these individual amino acids in the structure and stability of the dimer was accomplished by chemical modifications, multiple sequence alignment of legume Bowman-Birk inhibitors, and homology modeling. The state of association may also influence the physiological and functional role of these inhibitors.

Bowman-Birk inhibitors  $(BBIs)^1$  are small serine proteinase inhibitors found in the seeds of legumes in particular (1). Char-

acteristically, their molecular masses are in the range of 6–9 kDa. They are single polypeptides and comprise a binary arrangement of two sub-domains with a conserved array of seven disulfide bridges, which play a prominent role in the stabilization of their reactive site configuration (2, 3). These inhibitors interact, simultaneously and independently with two (not necessarily identical) molecules of proteinases (4) without any conformational change (5). The BBIs have two tandem homology regions comprising a consensus motif of three  $\beta$ -strands, each with a kinetically independent reactive site on the outermost exposed loop that adopts a common canonical conformation, similar to that of a productively bound substrate (1).

In addition to protease inhibitory activity, the anticarcinogenic activity and radioprotective activity of BBIs from legumes have been widely studied (6). Immune stimulating properties of these inhibitors have also been reported (7). The BBIs have been implicated to play a vital role in the arsenal defense mechanism that plants use to protect against insect predators and against environment hazards during germination and seedling growth.

Despite extensive studies on BBIs, only a few three-dimensional structures have been solved by x-ray or by NMR. These include the x-ray structure of PI-II from tracy bean (8), A-II from peanut (9), and soybean BBI (10, 11), which have been analyzed in the free form, and PsTI-IVb from pea seeds (5). The x-ray structure data of trypsin complexes with the BBIs from adzuki bean (12), mung bean (13), and soybean (14) are available. A three-dimensional model of black-eyed pea BBI-chymotrypsin complex has been constructed based on the homology of BBIs (15). The three-dimensional structure of a 16-kDa BBI from barley seeds at 1.9-Å resolution remains to be the highest resolution of a BBI to date (16). The x-ray structure of a novel and unique monofunctional 14-amino acid residue cyclic peptide, from sunflower seeds, complexed with trypsin, has exhibited both sequence and conformational similarity to the trypsin-reactive site loop of BBIs (17).

Horsegram (*Dolichos biflorus*), is a pulse crop native to Southeast Asia and tropical Africa. Four isoforms of BBIs, from horsegram seeds (18), have been isolated. The complete primary structure of the major isoform HGI-III has been determined (19). Three linear epitopes of the major inhibitor have been mapped, of which one contains the chymotrypsin inhibitory site (20). The role of disulfide linkages in maintaining the structural integrity of horsegram BBI was established by circular dichroism and fluorescence studies (21). Horsegram BBI followed the "two state" mode of unfolding and oxidative refolding of the BBI was possible only at very low inhibitor concen-

<sup>\*</sup> This work was supported in part by Grant SP/SO/A-70/97 from Department of Science and Technology, New Delhi, India. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>§</sup> Recipient of a Senior Research Fellowship from Council of Scientific
and Industrial Research (CSIR), New Delhi, India.

The amino acid sequence reported in this paper has been submitted to the Swiss Protein Database under Swiss-Prot accession no. Q9S9E3.

To whom correspondence should be addressed. Tel.: 91-821-251-5331; Fax: 91-821-251-7233; E-mail: lrg@cftri.com & lrgowda@yahoo.com.

<sup>&</sup>lt;sup>1</sup> The abbreviations used are: BBI, Bowman-Birk inhibitor; HGI, horsegram inhibitor; HGGI, horsegram germinated inhibitor; HPLC, high performance liquid chromatography; GuHCl, guanidium hydrochloride; BAPNA,  $\alpha$ -N-benzoyl-DL-arginine-p-nitroanilide HCl; DEPC,

diethylpyrocarbonate; TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone; MS, mass spectrometry; T, total acrylamide concentration; C, degree of cross-linking.

tration in a disulfide-thiol buffer (22). The three new isoforms that appear upon germination of horsegram seeds (23) are derived from the dormant seed inhibitors by a limited proteolysis during germination and not by de novo synthesis (24). The inhibitors of horsegram (HGIs) are single polypeptides with a molecular mass of 8.5 kDa. SDS-PAGE and analytical gel filtration indicate the molecular mass to be  $\sim 16$  kDa (18), suggesting that they exist as dimers in solution. Such self-association and anomalous behavior on SDS-PAGE resulting in a large overestimation of molecular mass has been reported for several legume BBIs (25-28). Many of the BBIs tend to undergo self-association to form homodimers or trimers or more complex oligomers (29). The three-dimensional model of the blackeyed pea BBI-chymotrypsin complex (15) and light scattering data (30) suggest that the inhibitor molecules are in continuous equilibrium between monomers and several forms of multimers. The data available on the protein-protein interactions, responsible for the self-association of BBIs, is sparse. In contrast to the dry seed inhibitors (HGIs), the inhibitors of germinated horsegram seeds (HGGIs), derived from the dry seed inhibitors, are single polypeptides of ~6.5–7.2 kDa and exist as monomers (24). In an attempt to understand and elucidate the structural features that contribute to the self-association of HGIs in solution, the primary structures of the HGGIs has been determined. The significant difference between the primary structures of inhibitors from the germinated seed (HG-GIs) and the inhibitor from the dormant seed (HGI-III) is the absence of the charged carboxyl-terminal tail and varied truncation at the amino terminus. This observation and the ability of HGI-III to self-associate and form dimers suggest that the structural elements responsible for this phenomenon occur at either the carboxyl and/or the amino terminus.

The dimeric crystal structure of the pea seed BBI, PsTI-IVb (5) reveals two monomers associated in a nearly perfect dimer that are mainly stabilized by an extensive hydrogen-bonded network, involving specific interactions between them, namely: (i) the guanidium group of  $\mathrm{Arg}^{23}$  of one monomer and the polar group of side chain of Glu<sup>68</sup> of second monomer and (ii) Lys<sup>16</sup> of one monomer and the dyad-related carboxyl group of Glu<sup>69</sup> of the other monomer. This observation and the finding that HGGIs, which lack the carboxyl-terminal Asp residues, lose the ability to form dimers suggest that these interactions play a unique role in the dimerization of BBIs. Based on these findings, chemical modification of the Lys/Arg residue, a comparative evaluation of the amino acid sequences of several BBIs that exist either as monomers or dimers and homology modeling of the dimer, we demonstrate the pivotal role of an interaction between Lys<sup>24</sup> (trypsin reactive site) and Asp<sup>76</sup> in HGI-III that characterizes the dimer formation. The effects of such dimerization on the functional aspects of the inhibitor are presented.

#### EXPERIMENTAL PROCEDURES

Materials—Horsegram (D. biflorus) seeds were obtained locally. DEAE-Sephacel and Sepharose-4B were obtained from Amersham Biosciences.  $\alpha$ -N-Benzoyl-DL-arginine-p-nitroanilide HCl (BAPNA), cyanogen bromide, diethylpyrocarbonate (DEPC), bovine pancreatic trypsin (2 × crystallized, type III, EC 3.4.21.4), L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin, endoproteinase Asp-N (EC 3.4.24.33), guanidine hydrochloride (GuHCl), 4-vinylpyridine, dithiothreitol, 1,2-dicyclohexanedione, and analytical gel filtration markers were procured from Sigma-Aldrich. Molecular weight markers for SDS-PAGE were from Bangalore Genei, Bangalore, India. All the other chemicals used were of highest purity.

*Purification of Isoinhibitors from Horsegram*—Dormant seed inhibitor (HGI-III) and horsegram germinated inhibitors (HGGI-I, -II, and -III) were purified as reported earlier (18, 24).

*Trypsin Inhibitory Assay*—The amidase activity of trypsin and its inhibition was assayed using the chromogenic substrate BAPNA as described earlier (24). One unit of trypsin 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, which reduces the enzyme activity by one unit.

*Polyacrylamide Gel Electrophoresis*—Denaturing SDS-PAGE (15% and 17.5% T with 2.7% C) of HGGI-III and HGI-III was performed according to the procedure of Laemmli (31). The gels were stained for protein with 0.1% Coomassie Brilliant Blue and destained.

Protein Sequencing—Reduction and alkylation of HGGIs with 4-vinyl pyridine was carried out as described for the dry seed inhibitors earlier (18). The pyridylethylated HGGIs were cleaved with TPCK-trypsin and endoproteinase Asp-N according to Aitken *et al.* (32). The pyridylethylated HGGIs were cleaved at the Met residue using a 50-fold molar excess of cyanogen bromide in formic acid over Met residues. Excess reagents were removed by repeated dilution and freeze-drying (33). The lyophilized digests were dissolved in 0.1% trifluoroacetic acid, and peptides were purified by HPLC using a C-18 reverse phase column (Phenomenex ODS column, 250 × 4.6 mm, 5  $\mu$ m) with 0.1% trifluoroacetic acid/CH<sub>3</sub>CN (7:3) linear gradient. The peptides were detected at 230 nm. The peak fractions were further rechromatographed using the same column and solvent system. The peptides were subjected to Edman analyses on an automated gas phase protein sequencer (Shimadzu PSQ-1).

Size-exclusion Chromatography—Size-exclusion measurements were performed using a BIOSEP-SEC-S 3000 (300  $\times$  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. The column was pre-equilibrated with the corresponding buffers prior to sample loading. The column was calibrated using a mixture of standard proteins, alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), cytochrome c (14.4 kDa), and HGGI-III (6.4 kDa).

SDS-PAGE analysis of  $Zn^{2+}$ -induced HGI-III Monomerization—The inhibitor was incubated in the presence of 1 mM ZnSO<sub>4</sub> for 1 h. The sample was then boiled with sample buffer for 5 min and separated by SDS-PAGE (17.5% T, 2.7% C) at pH 8.8. The separated proteins were then stained with 0.1% Coomassie Brilliant Blue.

Chemical Modification of HGI-III—Diethyl pyrocarbonate (DEPC) was used to modify His residues. DEPC solution was freshly prepared by dilution of the reagent in cold ethanol. The concentration of the stock was determined by reaction with 10 mM imidazole (34). For modification with DEPC, HGI-III was diluted in 0.1 M phosphate buffer, pH 7.25. At fixed time intervals, aliquots of DEPC were added to the mixture, and the formation of *N*-carbethoxyhistidine was monitored by the increase in absorbance at 240 nm, using a Shimadzu UV-1601 double beam spectrophotometer. The final concentration of DEPC ranged from 0 to 0.4 mM. The modified HGI-III was chromatographed on a BIOSEP-SEC-S 3000 column pre-equilibrated in phosphate buffer, pH 7.25.

Arginine residues of HGI-III were modified by the reviewed method of Smith (35) using 1,2-cyclohexanedione. HGI-III was dissolved in 0.2 M sodium borate buffer, pH 9.0 (0.1 mg/ml), and reacted with 0.15 M 1,2-cyclohexanedione at 35 °C for 2 h. The reaction mixture was acidified using 30% acetic acid and dialyzed against 10 mM acetic acid to remove excess reagents. The sample was concentrated and evaluated by SDS-PAGE.

Lysine residues of HGI-III were modified at  $25 \pm 2$  °C using citraconic anhydride (36). HGI-III (1 mg/ml) was dissolved in water and the pH adjusted to 8.0. Citraconic anhydride (1  $\mu$ l/mg of protein) was added to the solution and maintained at pH 8. The modified inhibitor was desalted and analyzed for the monomer-dimer status by SDS-PAGE as described earlier.

Modeling of HGI-III Dimer-The sequences of BBIs from leguminous plants were obtained from the NCBI protein sequence data base and aligned using the ClustalW multiple alignment algorithm (37). Of these, a crystal structure for the pea BBI inhibitor was available from the Protein Data Bank (PDB: 1PBI). The sequences of the monomeric HGGI-III and dimeric HGI-III were highly similar to that of the pea inhibitor, enabling building of their models by standard homology modeling techniques. Models of HGI-III and HGGI-III were built using the Biopolymer module in InsightII (Accelrys Inc.) and energy-minimized using DISCOVER. The quality of the structure was measured using PROCHECK (38). A single residue insertion at position 37, which forms part of a loop, was observed in HGI-III and HGGI-III, with respect to the crystal structure template. An analysis of related crystal structures in PDB using DALI (39) and Insight-II indicated that the structure of BBI proteinase inhibitor PI-II (1pi2) contained a similar insertion at the same position. The insertion in the loop is therefore modeled based on this structure. Visualization and analysis of the structures were carried out using Insight-II.

Thermal Stability Studies-The purified inhibitors were dissolved



FIG. 1. Summary of primary structure determination of germinated horsegram (*D. biflorus*) seed Bowman-Birk inhibitor, HGGI-III. *T*, *A*, and *CNBr* denote TPCK-trypsin, endoproteinase Asp-N, and CNBr fragments. respectively.

in water and incubated at 95  $\pm$  1 °C in a constant temperature water bath. Aliquots were removed at regular time intervals, immediately cooled on ice, and assayed for residual trypsin inhibitory activity as described earlier.

#### RESULTS

Primary Structure of the HGGIs and Comparison of Their Sequences—The complete amino acid sequences of HGGI-I, -II, and -III comprising of 66, 65, and 60 residues, respectively, were obtained by automated sequencing analysis of the proteins and peptides generated by enzymatic and chemical cleavage (Fig. 1, data shown only for HGGI-III). The molecular mass calculated on the basis of the sequence of HGGI-I, -II, and -III are 7109, 6993, and 6464, respectively. These results are in close agreement to the molecular mass determined by matrixassisted laser desorption ionization-MS, which were 7216.7, 7074.6, and 6493.5, respectively. The determined amino acid composition agrees with that of the deduced sequence (24).

The alignment of the amino acid sequences of the HGGIs with the sequence of HGI-III, the major isoinhibitor present in the dry seed is shown in Fig. 2. The sequences of the HGGIs are identical to HGI-III sequence, except for the truncation at both the amino and carboxyl termini of the sequence. The three inhibitors from germinated seeds (HGGIs) differ from each other only at the amino terminus. The absence of the charged tetrapeptide, -SHDD, at the carboxyl terminus is common to all the three HGGIs. The 14 half-cysteine residues are conserved in the HGGIs, as is the case of all legume BBIs sequenced thus far. The trypsin reactive site with Lys and chymotrypsin reactive site with Phe as the  $P_1$  residue in HGI-III remain unaltered in the sequence of the HGGIs. These results indicate that the three inhibitors of the germinated seed are *in situ* proteolytic products of the dry seed inhibitor HGI-III.

SDS-PAGE and Size-exclusion Chromatography—SDS-PAGE analysis of HGGI-III indicates it is a single polypeptide with a molecular mass of ~6.5 kDa (Fig. 3). HGGI-I and HGGI-II also move as single polypeptides of ~7.0-kDa molecular mass on SDS-PAGE (24). In contrast, HGI-III moves as a single polypeptide of ~16.0-kDa molecular mass (Fig. 3, *lane C*). The exact molecular mass of HGI-III as determined by electrospray mass spectrometry (18) and by sequence (19) is ~8.0 kDa. These results suggest that the HGI-III in solution undergoes self-association to form a dimer. The reduced and alkylated HGI-III is a polypeptide of ~8.0 kDa (18).

The monomer/dimer status of HGI-III and HGGI-III was further evaluated by size-exclusion HPLC on a BIOSEP-SEC-S 3000 column using 0.25  $\times$  Tris-HCl, pH 7.25. HGI-III was well separated from HGGI-III (Fig. 4, *A* and *B*). HGGI-III eluted later with a retention time of 22.92 min corresponding to a molecular mass of 6.5 kDa. HGI-III elutes at 20.26 min, which corresponds to a molecular mass of ~16.0 kDa. These results provide further evidence that HGI-III in solution associates to form a dimer.







FIG. 3. **SDS-PAGE (15% T, 2.7% C) of HGI-III and HGGI-III.** Lane A, low molecular weight markers; lane B, HGGI-III; lane C, HGI-III; and lane D, high molecular weight markers.



FIG. 4. Size-exclusion chromatography of HGI-III and HGGI-III. The samples were dissolved in different buffers and loaded on to a BIOSEP-SEC-S 3000 column pre-equilibrated with respective buffers and eluted at 0.5 ml/min. *A*, HGI-III (pH 7.25); *B*, HGGI-III (pH 7.25); *C*, HGI-III (1 mM ZnSO<sub>4</sub>, pH 7.25); *D*, HGI-III (1 mM ZnCl<sub>2</sub>, pH 7.25); and *E*, HGI-III (1 mM ZnCl<sub>2</sub>, pH 6.5).

Rationale for Chemical Modification—A closer evaluation of the sequences shows that the most significant difference between HGI-III and HGGI-III, is the physiological deletion of the peptide -DHHQSTDEPSES and the tetrapeptide -SHDD at the amino and carboxyl termini, respectively. Hence, either the depleted amino and/or the carboxyl termini are involved in the self-association of HGI-III. HGI-I, yet another of the isoinhibitors present in the dry seed of horsegram, although truncated



FIG. 5. **SDS-PAGE** (15% **T**, 2.7% **C**) of chemically modified HGI-III. *Lane A*, high molecular weight markers; *lane B*, unmodified HGI-III; *Lane C*, HGI-III modified with dicyclohexanedione; and *lane D*, HGI-III modified with citraconic anhydride.

at the amino terminus, exists as a dimer in solution (18). These observations implicate that the residues of the deleted tetrapeptide play a vital role in the dimerization of HGI-III. PsTI-IVb, a BBI from Winter pea seeds, has been crystallized as a nearly perfect 2-fold symmetric dimer in the asymmetric subunit, which includes its carboxyl-terminal segment (5). The carboxyl-terminal tail from residues 68 to 70 (EEV), which constitutes an extended  $\beta$ -strand, makes no contact with its own subunit yet is held by interactions with the other subunit. Two specific interactions that have been discerned between the two subunits are (a) a hydrogen bond between the guanidium group of Arg<sup>23</sup> of one subunit and the polar group of the side chain of  $\operatorname{Glu}^{68}$  and (b) an ion pair between  $\operatorname{Lys}^{16}$  of one subunit and the dyad-related carboxyl group of Glu<sup>69</sup> of the other subunit. This observation, together with the fact that the deleted tetrapeptide contained Asp residues, suggests that such interactions could well be the premise to self-association in HGI-III. The effect of chemically modifying Arg and Lys residues of HGI-III has been studied. Chemical modification of these residues may disrupt such a subunit interaction in HGI-III leading to the formation of monomers.

Chemical Modification of Arg and Lys Residues of HGI-III-The Lys residues of HGI-III were chemically modified using citraconic anhydride. Citraconylation resulted in the acetylation of the free  $\epsilon$ -amino group of Lys. SDS-PAGE of the modified HGI-III revealed an increased relative mobility compared with the unmodified inhibitor (Fig. 5, *lane* D). The molecular mass, calculated on the basis of the relative mobilities of a set of standard proteins, was ~8.5 kDa, corresponding to that of a monomer. The conversion of the dimer form of the HGI-III to monomer by citraconylation suggests that a Lys residue is involved in the self-association of HGI-III. The guanidium group of Arg residues were modified using 1,2-cyclohexanedione, resulting in a heterocyclic condensation product between the guanidium group of Arg and the carbonyl of 1,2-cyclohexanedione. SDS-PAGE (15% T, 2.7% C) analysis at pH 8.8 led to the understanding that the modification had no effect on the relative mobility of HGI-III (Fig. 5, lane C). The relative mobilities of both the unmodified (Fig. 5, lane A) and the modified inhibitor were identical (Fig. 5, lanes B and C). These results ruled out the essentiality of an interaction involving Arg for self-association of HGI-III.

Effect of  $Zn^{2+}$  on the Self-association of HGI-III—The changes in the monomer-dimer status of HGI-III, in the presence of  $Zn^{2+}$ , was followed by size-exclusion chromatography on a BIOSEP-SEC-S 3000 column, and the results are summarized in Fig. 4. HGI-III elutes as a single symmetrical peak with a retention time of 20.26 min (Fig. 4A). In the presence of either 1 mM ZnCl<sub>2</sub> or ZnSO<sub>4</sub> at pH 7.25, HGI-III eluted at 22.65 min (Fig. 4, *C* and *D*). This increased retention time pointed to a significantly reduced molecular weight,



FIG. 6. SDS-PAGE (17.5% T. 2.7% C) of  $Zn^{2+}$ -induced monomerization of HGI-III. Lane A, molecular weight markers; lanes B and C, HGI-III incubated in the presence and absence of 1 mM ZnSO<sub>4</sub> for 1 h, respectively.

calculated to be ~8.5 kDa. In contrast, when similar experiments were performed in buffer of pH 6.5 containing 1 mM  $\text{ZnSO}_4$  or  $\text{ZnCl}_2$ , HGI-III eluted at 20.27 min (Fig. 4*E*). These results suggested that HGI-III can be converted to its monomeric form in the presence of  $\text{Zn}^{2+}$  at neutral pH. This dissociation of HGI-III, in the presence of  $\text{Zn}^{2+}$ , was also followed on SDS-PAGE (Fig. 6). The treatment of HGI-III with  $\text{Zn}^{2+}$  increased its relative mobility. The molecular mass of the  $\text{Zn}^{2+}$ -modified HGI-III was ~8.4 kDa. These results concurred with those of size-exclusion chromatography.

The dissociation of HGI-III to monomers in the presence  $\mathrm{Zn}^{2+}$  at pH 7.25 and the absence of monomer formation in the presence of  $Zn^{2+}$  at pH 6.5 and below (results not shown) suggested the probable involvement of a His residue. DEPC was used to modify HGI-III to evaluate the role of His in the self-association phenomenon. DEPC-modified HGI-III eluted at 20.26 min, which corresponds to the dimeric form of HGI-III, ruling out the involvement of His in the monomer interaction. His residues in proteins are found to bridge both a Zn<sup>2+</sup> ion and a carboxylate side chain of nearby aspartates (sometimes glutamate residues) referred to as an indirect carboxylate-metal co-ordinate (40). The conversion of HGI-III to its monomer in the presence of Zn<sup>2+</sup> at pH 7.25 probably occurs through this indirect metal co-ordination, involving Asp (Asp<sup>75/76</sup>) residues at the carboxyl terminus of HGI-III. These results indirectly provide evidence that the Asp residues at the carboxyl terminus play a vital role in the dimerization of HGI-III. These results, together with the observation that a modification of Lys and not Arg causes the formation of monomers, support the inference that the interaction between Lys of one monomer and Asp of the other monomer is responsible for the self-association of HGI-III. Pair-wise alignment of the PsTI-IVb with HGI-III indicates that this interaction in the HGI-III dimer should occur between Lys<sup>24</sup> at the trypsin reactive site and Asp<sup>76</sup> at the carboxyl terminus.

Comparative Evaluation of BBI Amino Acid Sequences in Relation to Monomer-Dimer Status-The BBI sequences of legumes were aligned to further evaluate the probability of such a unique interaction, between the first reactive site residue and Asp/Glu at the carboxyl terminus, dictating the monomerdimer status of BBIs in solution. The alignment of BBI sequences of dicots obtained from Swiss-Prot (Release 42, October 2003) and those reported in the literature is shown in Fig. 7. The program MULTALIN (41) was used to align the sequences. Some dicot BBIs exist as monomers in solution despite the close similarity between their sequences, whereas others self-associate to form dimers. A closer look at the first reactive site residue and the carboxyl-terminal sequences in relation to the monomer-dimer status reveals the following: (i) If the first reactive site in the dicot BBIs is Lys/Arg at the P<sub>1</sub> position and the carboxyl terminus contains an Asp/Glu resi-

	1 10	20	30	40 43
UCTT_TV	venuecen	necevoren	СИСТОСИРООС	
CDT_C_TT	MEI NI EVENUECENI	DESSKPULD		
PVT_3T	GHR7HSR7P	SZSSKPCCB	ICVCTASTPP7C	VCTRTPIN
CRI-TI		SZSSKICCB.	TEVETASTEPOE	NCTRTPI R
MRT	SHDEP	SESSEPTCED.	SCULTKSKDDUC	HCANTPIN
DE-4	DODHSDOFPRE	SESSEPTICS	SC-CTRSRPPOC	OCTOVELN
T-TT ID	DHSDDE	SESSKPCCD	FCKCTKSEPPOC	OCUNTRI F
CLTI-II	DHSDDE	SESSKPCCD	ECKCTKSEPPOC	OCYDTRLE
PVI-3II	SGHRHESZBSTBZA	SZSSKPCCB	ICACTKSIPPOC	RCSBLRLN
BTCI	SGHHZBSTBZA	SZSSKPCCR	ZCACTKSIPPZC	RCSZYRLN
LBI	SGHHEHSTDZP	SZSSKPCCB	ICACTKSIPPQC	RCTDLRLD
HGI-III	DHHQSTDEP	SESSKPCCD	DCACTKSIPPOC	RCTDYRLN
MAI-DE-3	DHHHSTDEP	SESSKPCCD	ECACTKSIPPQC	RCTDYRLN
HGGI-I	DEPS	SESSKPCCD	QCACTKSIPPQC	RCTDYRLN
HGGI-II	EPS	SESSKPCCD	QCACTKSIPPQC	RCTDYRLN
BBI	DI	DESSKPCCD	QCACTKSNPPQC	RCSDMRLN
HGGI-III		SKPCCD	QCACTKSIPPQC	RCTDYRLN
IcII2		SSKheaccdi	RCACTKSIPPQC	HCACIRLN
TaTI		SSKheaccdi	RCACTKSIPPQC	HCADIRLN
PI-II		DEYSKPCCD	CHCTRSHPPQC	SCED-RIN
MSTI	T	KSTTTSCCD	CPCTRSIPPQC	QCTDYR-E
HII		TTACCN	CPCTRSIPPQC	RCTDIG-E
FBI	G	DDYKSACCD	ICLCTKSEPPTC	RCYDYG-E
PSI1-IVb	G	UDVKSHCCD	ICLUIKSNPPIC	RCYDYG-E
Consensus	••••••	s.sskptC#	.C.CIKSIPPqC	•C•#•L1#
	44 53	63	73	83 86
	44 53	63 +	73	83 86 +1
HSTI-IV	44 53 I	63 HPGQCRCLD	73 ITDFCYKPCKSS	83 86 1 DEDDD
HSTI-IV SBI-C-II	44 53 I	63 HPGQCRCLD HPGQCRCLD	73 TTDFCYKPCKSS TTDFCYKPCKSS	83 86 1 DEDDD DEDDD
HSTI-IV SBI-C-II PVI-3I	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI	63 HPGQCRCLD HPGQCRCLD HPGZCRCLB	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI	63 HPGQCRCLD HPGQCRCLD HPGZCRCLB HPGKCRCLB	73 ITDFCYKPCKSS ITDFCYKPCKSS ITBYCYKSCKSB ITBYCYKSCKSD	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II HBI	44 53 SCHSRCSI SCHSRCDRCRCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCICTRSI	63 HPGQCRCLD HPGQCRCLD HPGZCRCLB HPGKCRCLB HPGKCRCLD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDDFCYKPCESH	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II HBI DE-4	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTTSI SCHSACKSCHCTTSI	63 HPGQCRCLD HPGQCRCLD HPGZCRCLB HPGKCRCLB HPGKCRCLD DPGHCSCLD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDDFCYKPCESH YTDFCYKPCKSS	83 86 1 DED DD DED DD SGZ BB SGZ BB SGZ BB DKD GDD DZZ
HSTI-IV SBI-C-II PVI-3I GBI-II MBI DE-4 CLII-I	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTFSI SCHSACKSCHCTFSI SCHSACKLCLCRLSI	63 HPGQCRCLD HPGQCRCLD HPGZCRCLB HPGKCRCLB HPGKCRCLD DPGHCSCLD FPAKCRCVD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDDFCYKPCESH YTDFCYKPCKSS TTDFCYKPCKSS	83 86 1 DEDDD DEDDD SGZBB SGZBB SGZBB DKD GDDDZZ GGDED GDDED
HSTI-IV SBI-C-II PVI-3I GBI-II MBI DE-4 CLTI-I CLTI-I	44 53 SCHSAC-SS SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKLCLCALSI	63 HPGQCRCLD HPGQCRCLD HPGQCRCLB HPGKCRCLB HPGKCRCLD DPGHCSCLD FPAKCRCVD FPAKCRCVD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS	83 86 1 DED DD SGZ BB SGZ BB SGZ BB DKD GDD DZZ GGD EDD GGD EDD UCD DDUU
WSTI-IV SBI-C-II PVI-3I GBI-II MBI DE-4 CLTI-I CLTI-II PVI-3II	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKLCLCALSI	63 HPGQCRCLD HPGQCRCLD HPGKCRCLB HPGKCRCLD DPGHCSCLD DPGHCSCLD FPAKCRCVD FPAKCRCVD IPAQCICTD IPAQCICTD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS TNFCYEPCKSS TNFCYEPCKSS	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLII-I CLII-II PVI-3II BTCI	44 53 SCHSRCSI SCHSRCDRCRCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKLCLCRLSI SCHSRCKLCLCRLSI SCHSRCKLCLCRLSI SCHSRCKSCRCTFS: SCHSRCKSCRCTFS:	63 HPGQCRCLD HPGQCRCLD HPGCRCLB HPGKCRCLB HPGKCRCLD DPGHCSCLD DPGHCSCLD FPAKCRCVD FPAKCRCVD IPAQCICTD IPAQCICTD IPAQCICCD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDFCYKPCKSG TDFCYKPCKSG TTDFCYKPCKSG TNFCYEPCKSS BBFCYKPCKSS BBFCYKPCKSS	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II MBI DE-4 CLTI-I CLTI-II CLTI-II PVI-3II BTCI LBI UCT_II	44 53 SCHSACCHCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKSCACTFSI SCHSACKSCACTFSI SCHSACKSCACTFSI	63 HPGQCRCLD HPGQCRCLB HPGKCRCLB HPGKCRCLD DPGHCSCLD DPGHCSCLD IPAGKCRCVD IPAQCICTD IPAQCICTD IPAQCYCBB IPAQCYCBB	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TDFCYKPCKSS TDFCYKPCKSS TTDFCYKPCKSS TDFCYKPCKSS EBBFCYKPCKSS EBBFCYEPCKSS	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-I CLTI-II PVI-3II BTCI LBI HGI-III HGI-IF-2	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKSCMCTFS: SCHSACKSCMCTFS: SCHSACKSCMCTFS:	63 MPGQCRCLD MPGQCRCLD MPGKCRCLD MPGKCRCLB MPGKCRCLD DPGMCSCLD FPAKCRCVD FPAKCRCVD FPAKCRCVD IPAQCCCD IPAQCCCBB IPAQCVCVD IPAQCVCVD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDBYCYKSCKSD TDDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS EBBFCYKPCKSS EBBFCYEPCKSS KDFCYAPCKSS	83 86 +1 DEDDD DEDDD SG2BB SG2BB SG2BB GDD22 GGDEDD GGDEDD HGPBBNN HSBBBBHN HSBBBBHN HSDDNNN HDD
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-I CLTI-II PVI-3II BTCI LBI HGI-III HHI-DE-3 HGE-I	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTFSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKSCACTFSI SCHSACKSCICTLSI SCHSACKSCICTLSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI	63 HPGQCRCLD HPGQCRCLD HPGKCRCLB HPGKCRCLB HPGKCRCLD HPGKCRCLD HPGKCRCVD FPAKCRCVD FPAKCRCVD IPAQCCCD IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDDFCYKPCKSS TTDFCYKPCKSS EBBFCYKPCKSS EBFCYKPCKSS EKDFCYAPCKSS HKDFCYAPCKSS HKDFCYAPCKSS	83 86 +1 DEDDD DEDDD SG2BB SG2BB SG2BB OKD GDDD22 GGDEDD GGDEDD HGP BBNN HSBBBBHN HSBBBBHN HSDDNNN HDD
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-I CLTI-II PVI-3II BTCI LBI HGI-III HGI-III HGG-III	44 53 SCHSRCSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKLCLCRLSI SCHSRCKLCLCRLSI SCHSRCKSCICTLS: SCHSRCKSCICTLS: SCHSRCSSCVCTFS: SCHSRCSSCVCTFS: SCHSRCSSCVCTFS:	63 MPGQCRCLD HPGQCRCLD HPGCRCLB HPGKCRCLD DPGMCSCLD PFAKCRCVD FPAKCRCVD IPAQCICTD IPAQCICTD IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDBYCYKSCKSD TDDFCYKPCKSG TTDFCYKPCKSG TTDFCYKPCKSS IBDFCYEPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKS I	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II BE-4 CLTI-I CLTI-I CLTI-II PVI-3II BTCI LBI HGI-III HGI-III HGGI-II HGGI-II BBI	44 53 SCHSRCSI SCHSRCDRCRCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTRSI SCHSRCKSCHCTFSI SCHSRCKSCHCTFSI SCHSRCSSCVCTFSI SCHSRCSSCVCTFSI SCHSRCSSCVCTFSI SCHSRCSSCVCTFSI SCHSRCSSCVCTFSI SCHSRCSSCVCTFSI	63 HPGQCRCLD HPGQCRCLD HPGCRCLB HPGKCRCLB HPGKCRCLD DPGHCSCLD IPGKCRCVD IPAQCICTD IPAQCICTD IPAQCVCVDI	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDFCYKPCKSG TDFCYKPCKSG TTDFCYKPCKSG TTDFCYKPCKSS INFCYEPCKSS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II MBI DE-4 CLII-I CLII-II CLII-II BTCI BTCI HGI-III HGGI-II HGGI-II HGGI-II HGGI-III	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTFSI SCHSACKSCMCTFSI SCHSACKSCMCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI	63 MPGQCRCLD' MPGQCRCLD' MPGKCRCLD' MPGKCRCLB' MPGKCRCLD' MPGKCRCLD' IPGKCRCLD' IPGKCRCLD' IPAQCCCD' IPAQCVCVDI IPA	73 TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TTDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS ITDFCYKPCKSS IKDFCYAPCKSS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS IKDFCYAPCKS	83 86 +1 DEDDD DEDDD SG2BB SG2BB SG2BB GDD22 GGDEDD GGDEDD HGP BBNN HSBBBBHN HSBBBBHN HSDDDNNN HDD EDDKEN
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-I CLTI-II CLTI-II BTCI HGI-III HGGI-II HGGI-II HGGI-III LCII2	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTFSI SCHSACKSCHCTFSI SCHSACKSCACTFSI SCHSACKSCUCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI	63 MPGQCRCLD MPGQCRCLD MPGCRCLB MPGKCRCLB MPGKCRCLD DPGKCSCLD FPAKCRCVD FPAKCRCVD IPAQCCCVD IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDDFCYKPCKSG TTDFCYKPCKSG TTDFCYKPCKSG INNFCYEPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKS IKDF	83 86 +1 DEDDD DEDDD SG2BB SG2BB SG2BB GDDDZZ GGDEDD GGDEDD HGP BBNN HSB BBNN HSD BNNN HDD EDDKEN
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-I CLTI-II CLTI-II PVI-3II BTCI LBI HGI-III HGGI-II HGGI-II HGGI-III HGGI-III TCTI2 TAT	44 53 SCHSAC-SSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKSCHCTFSI SCHSACKSCHCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI	63 HPGQCRCLD HPGQCRCLD HPGKCRCLB HPGKCRCLB HPGKCRCLD DPGMCSCLD FPAKCRCVD FPAKCRCVD IPAQCICTD IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCCCVDI IPAQCCCVDI IPAQCRCFD IPAQCRCFD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFCYAPCKS TCDFC	83 86 +1 DEDDD DEDDD SG2BB SG2BB SG2BB OKD GDDDZZ GGDED GGDEDD HGP BBNN HSBBBHN HSBBBHN HSD BBNN HDD EDDKEN
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-I CLTI-II PVI-3II BTCI BTCI HGI-III HGI-III HGGI-II HGGI-II BBI HGGI-III TcTI2 TaTI PI-II	44 53 SCHSACSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTFSI SCHSACKLCLCALSI SCHSACKSCICTLS: SCHSACKSCICTLS: SCHSACKSCICTLS: SCHSACSSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSCVC	63 HPGQCRCLD HPGQCRCLD HPGCRCLB HPGKCRCLD DPGMCSCLD PFAKCRCVD FPAKCRCVD IPAQCICTD IPAQCICTD IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCCCVDI IPAQCRCFD IPAQCRCFD IPAQCRCFD IPAGCRCFD IPAGCRCFD	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDBCYKSCKSD TDDFCYKPCKSG TTDFCYKPCKSG TDFCYKPCKSS IBDFCYEPCKSS IKDFCYAPCKSS IKDFCYAPCKSS IKDFCYAPCKS IKDFCYAPCKS IKDFCYCKPCSG IRDFCYKPCSG IRDFCYKPCSG	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II MBI DE-4 CLTI-I CLTI-II CLTI-II PVI-3II BTCI BBI HGGI-III HGGI-II HGGI-II HGGI-III TcTI2 TaTI PI-II HSTI	44 53 SCHSACSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTFSI SCHSACKSCMCTFSI SCHSACKSCMCTFSI SCHSACKSCMCTFSI SCHSACSSCVCTFSI SCHSACSCVCTFS	63 MPGQCRCLD' MPGQCRCLD' MPGCRCLD' MPGKCRCLB' MPGKCRCLB' MPGKCRCLD' IPGKCRCLD' IPAGKCRCVD' IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCRCFD' IPAKCRCFD' IPPQCRCVD'	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDBYCYKSCKSD TDDFCYKPCKSG TTDFCYKPCKSG TTDFCYKPCKSG TNFCYEPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSG TDFCYKPCSG TNDFCYKPCSG TNDFCYKPCSG TNDFCYKPCSS	83 86 
HSTI-IV SBI-C-II PVI-3I GBI-II MBI DE-4 CLTI-II CLTI-II CLTI-II BTCI BTCI BBI HGGI-III HGGI-II HGGI-III BBI HGGI-III TCI12 TaTI PI-III HII	44 53 SCHSACSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTRSI SCHSACKSCMCTFSI SCHSACKSCMCTFSI SCHSACKSCACTFSI SCHSACKSCMCTFSI SCHSACSSCVCTFSI SCHSACSCVCTFSI S	63 MPGQCRCLD' MPGQCRCLD' MPGKCRCLD' MPGKCRCLB' MPGKCRCLCB' MPGKCRCVD' FPAKCRCVD' FPAKCRCVD' IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCCCVDI IPAQCCCVDI IPAQCCCVDI IPAQCCCVDI IPAQCCCVDI IPAQCCCCDI IPAQCRCCD' IPPQCRCCD'	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDBYCYKSCKSD TDDFCYKPCKSS TTDFCYKPCKSS TTDFCYKPCKSS TNFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKS TKDFCYAPCKS TKDFCYAPCKS TCDFCYKPCSG TDFCYKPCSG TDFCYKPCSG TDFCYKPCKS TDFCYKPCKS TDFCYKPCKS	83 86 +1 DEDDD DEDDD SG2BB SG2BB SG2BB GDD22 GGDEDD GGDEDD GGDEDD HGP BBNN HSBBBBHN HSBBBBHN HSDDDNNN HDD EDDKEN
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-II CLTI-II CLTI-II BTCI BTCI HGI-III HGI-III HGGI-II HGGI-II TCTI2 TATI PI-II HSTI HII HGII FEI	44 53 SCHSAC-SSCHCTRSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTFSI SCHSACKCLCALSI SCHSACKSCLCLCALSI SCHSACKSCLCLCALSI SCHSACKSCLCTFSI SCHSACKSCLCTFSI SCHSACSSCVCTFSI SCHSACSCVCTFSI SCHSACSCVCTFSI SCHSACSSCVCTFSI SCHSACSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSC	63 MPGQCRCLD MPGQCRCLB MPGKCRCLB MPGKCRCLB MPGKCRCLD MPGKCRCVD FPAKCRCVD FPAKCRCVD IPAQCCCVD IPAQCVCVDI	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSB TTBYCYKSCKSD TDBCYKPCKSG TTDFCYKPCKSG TTDFCYKPCKSG TNFCYEPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKS TKDFCYAPCKS TKDFCYAPCKS TKDFCYAPCKS TKDFCYAPCKS TKDFCYAPCKS TKDFCYAPCKS TKDFCYCKPCSG TKDFCYKPCSG TKDFCYPSCR TTNFCYPSCR TTNFCYPSCR	83 86 +1 DEDDD DEDDD SG2BB SG2BB SG2BB OKD GGDEDC GGDEDD HGP BBNN HSB BBHN HSB BBHN HSD DNNN HDD EDDKEN
HSTI-IV SBI-C-II PVI-3I GBI-II DE-4 CLTI-I CLTI-II PVI-3II BTCI LBI HGI-III HGGI-II HGGI-II HGGI-III TCTI2 TATI PI-II HSTI HSTI FBI PSTI-IV	44 53 SCHSACSI SCHSACDRCACTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKSCHCTRSI SCHSACKLCLCALSI SCHSACKLCLCALSI SCHSACKSCACTFSI SCHSACKSCACTFSI SCHSACKSCACTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSCCCTSI SCHSACKSCACTRSI KCHBACKSCLCTKSI TCHSACKSCLCATSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACSSCVCTFSI SCHSACKSCACTRSI KCHBACKSCLCTKSI TCHSACLSCLCATSI	63 HPGQCRCLD HPGQCRCLD HPGCRCLB HPGKCRCLB HPGKCRCLD DPGMCSCLD FPAKCRCVD FPAKCRCVD IPAQCICTD IPAQCICTD IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCVCVDI IPAQCCCVDI IPAQCCCVDI IPAQCRCCDI IPPQCRCCDI IPPQCRCCDI IPPQCRCCDI IPPQCCCDI IPPQCCCDI IPPQCCCDI IPPQCCCDI IPPQCCCDI IPPQCCCDI IPPQCCCDI IPPCCCDI IPPCCCCDI IPPCCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPCCDI IPPC	73 TTDFCYKPCKSS TTDFCYKPCKSS TTBYCYKSCKSD TDBYCYKSCKSD TDFCYKPCKSG TTDFCYKPCKSG TTDFCYKPCKSG TDFCYKPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TKDFCYAPCKSS TDFCYKPCSG TNDFCYKPCSG TNDFCYKPCSG TNDFCYKSCHN TQKFCYKSCHN	83 86 +1 DEDDD DEDDD SGZBB SGZBB SGZBB GDDZZ GGDED GGDEDD HGP BBNN HSBBBHN HSBBBHN HSDDNNN HDD EDDKEN

FIG. 7. Multiple alignment of legume Bowman-Birk inhibitor sequences. The sequence alignment was performed using MUL-TALIN. The *arrows* are between the  $P_1$  and  $P_1'$  residues of the reactive site. Abbreviations used are as described in Table I.

due, the BBI exhibits self-association to form dimers, namely HGI-III, BTCI, PsTI-IVb, CLTI-I, CLTI-II, and BBI (Fig. 7 and Table I). (ii) If the P<sub>1</sub> position of the first reactive site is occupied by Ala in place of Arg/Lys as in the elastase inhibitors, SBI-C-II, WSTI-IV, and PVI-3I, the BBI exists as a monomer in solution. (iii) If the first reactive site is occupied by Lys/Arg at the P<sub>1</sub> position but lack Asp/Glu residues at the carboxyl-terminal end, the BBI exists as a monomer in solution, *e.g.* HGGI-I, -II, and -III, WII, MSTI, TcTI2, and TaTI, and FBI. (iv) BBIs that have the Arg<sup>23</sup> involved replaced by the other residues such as His or Gln or Ser yet exist as dimers or are self-associated in solution (DE-4, CLTI-I, and CLTI-II).

These observations strongly support the fact that the unique interaction between the Arg/Lys at  $P_1$  of the first reactive site of one subunit and Asp/Glu at the carboxyl-terminal of the second subunit is responsible for the observed self-association in BBIs. In HGI-III, this interaction is between Lys<sup>24</sup> of one subunit and Asp<sup>76</sup> of the second subunit. The loss of Asp<sup>76</sup> of the HGI-III, during germination, to form HGGI-III, disrupts this interaction, which leads to HGGIs being monomers.

Homology Modeling of the HGI-III Dimer—The sequences of HGI-III and HGGI-III are identical to each other except at the

amino and carboxyl termini, which clearly indicates that the same subunit structure is exhibited by both the molecules and the changes between them are restricted to the two termini. The models (Fig. 8) reveal that individual subunits belong to the knottins fold (57), formed by very few secondary structural elements, stabilized primarily by the seven disulfide bonds, similar to that in the template. HGI-III has an extension of 12 residues at the amino terminus and an extension of four residues at the carboxyl terminus, as compared with that in HGGI-III. The 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 (Fig. 8, inset). This will not influence the stability of the dimer. The carboxyl termini, however, are located at the dimer interface and play an important role in the dimer stabilization. The interactions made by the carboxyl-terminal segment is in fact present in 2-fold at the interface, due to the contributions from both subunits. In particular, the carboxyl-terminal Asp from one subunit  $(Asp^{A76} \text{ and } Asp^{B76})$  forms salt bridges with Arg from the other subunits  $(Arg^{B31} \text{ and } Arg^{A31}, respectively)$ as illustrated in Fig. 8. The two Arg and Asp in turn are held in appropriate orientations for interaction by a further network of hydrogen bonds with their neighborhood, involving Gln<sup>19</sup> and Lys<sup>24</sup> in both the subunits, respectively (Fig. 8). These observations suggest that only when Lys<sup>24</sup> of one subunit is hydrogen-bonded to Asp<sup>76</sup> of the second subunit, Arg<sup>31</sup> (A subunit) and Asp<sup>76</sup> (B subunit) are in juxtaposition to form the salt bridge resulting in the dimer. In the absence of this critical Lys<sup>24</sup>-Asp<sup>76</sup> hydrogen bond, Arg<sup>31</sup> and Asp<sup>76</sup> are not in a favored orientation to form the salt bridge. The situation is similar in the pea inhibitor and in fact in all the BBI structures known, indicating that the carboxyl-terminal segment is the determining factor in dimer formation. In HGGI-III, Asp<sup>76</sup> is not present due to the truncation of the polypeptide chain during germination (Fig. 1) and therefore cannot form the dimer, correlating well with the observation of a monomer from biochemical studies (24).

Thermal Stability of HGI- and HGGI-III—To understand the effect of the monomer/dimer status upon the physiological properties the thermal stability of the inhibitors was evaluated. Fig. 9 represents the changes in trypsin inhibitory property at 95 °C. For HGI-III, relatively little or no changes in the trypsin inhibitory activity were observed up to 6 h (Fig. 9). Although the stability of HGGI-III was comparable to HGI-III up to 3 h, HGGI-III lost 50% of its activity after 6-h incubation at 95 °C. The decrease in the thermal stability led to the suggestion that the dimeric form of the inhibitor is more stable than the monomer.

#### DISCUSSION

Legume seeds are well known for the large protein reserves in their cotyledons and for their levels of protein-proteinase inhibitors active against mammalian pancreatic proteinases. The BBIs present in the resting horsegram seed rapidly disappeared with the concomitant appearance of new active species during germination. To correlate the structure and stability of these interrelated variant forms at the molecular level, the amino acid sequence of the multiple inhibitors during germination was established.

HGI-III, the major BBI of horsegram seed is a very stable globular protein consisting of a single polypeptide chain of 76 amino acid residues (19). The primary structure of the three HGGIs (Fig. 2) derived from germinated seeds differed from the HGI-III at the amino and carboxyl termini. From the comparison of the sequences, it is clear that the cleavage occurs between Thr<sup>6</sup>-Asp<sup>7</sup>, Asp<sup>7</sup>-Asp<sup>8</sup>, and Ser<sup>12</sup>-Ser<sup>13</sup> at the amino terminus in the conversion of HGI-III to HGGI-I, HGGI-II, and

### Self-association of Bowman-Birk Inhibitors

Common name (botanical name)	Abbreviation	Monomer/dimer status	Enzyme inhibited at first reactive site	Reference
Wild soja (Glycine soja)	WSTI-IV	Monomer	Elastase	42
Bushbean (Phaseolus vulagris var. nanus)	PVI-3I	Monomer	Elastase	43
	PVI-3II	$\mathrm{NR}^a$	Trypsin	
Torresea cearensis (Amburana cearensis)	TcTI2	Monomer	Trypsin	44
	TaTI	Monomer	Trypsin	45
Black eyed pea (Vigna unguiculata (L))	BTCI	Dimer	Trypsin	46
Tracy bean (Glycine max)	PI-II	NR	Trypsin	47
Snail medic seeds (Medicago scutellata)	MSTI	Monomer	Trypsin	48
Alfa alfa leaves (Medicago sativa)	WII	Monomer	Trypsin	49
Faba beans (Vicia faba L)	FBI	Monomer	Trypsin	50
Pea seeds	PsTI-IVb	Dimer	Trypsin	5
Horsegram (Dolichos biflorus)	HGI-III	Dimer	Trypsin	19
	HGGI-I	Monomer	Trypsin	24
	HGGI-II	Monomer	Trypsin	24
	HGGI-III	Monomer	Trypsin	24
Garden bean(Phaseolus vulgaris)	GBI-II	NR	Elastase	51
Soybean (Glycine max)	BBI	Dimer	Trypsin	2
	SBI-C-II	Monomer	Elastase	52
Mung bean (Vigna radiata or Phaseolus aureus)	MBI	NR	Trypsin	53
Apple leaf (Lonchocarpus capassa)	DE-4	Dimer	Trypsin	54
Macrotyloma axillare	MAI-DE-3	NR	Trypsin	55
Lima bean (Phaseolus lunatus)	LBI	Dimer	Trypsin	56
Canavalia lineata	CLTI-I	Dimer	Trypsin	27
	CLTI-II	Dimer	Trypsin	

TABLE I								
Comparison	$of \ self\ association$	status of BBIs	s of leguminous	seeds				

<sup>a</sup> NR, not reported.

HGGI-III, respectively (Fig. 2). The loss of the tetrapeptide -SHDD during germination indicates a cleavage at  $\text{Ser}^{72}$ - $\text{Ser}^{73}$ of HGI-III. The cleavage is due to the highly specific action of an *in situ* protease. A metalloprotease, purified from germinated horsegram, converted HGI-III to HGGIs.<sup>2</sup> The three inhibitors that appear on germination are not the products of stored mRNA but derived from the dry seed inhibitor by a limited proteolysis at both amino and carboxyl termini. Sequence determinations of two electrophoretically distinct inhibitors (MBI-E and MBI-F), which appear upon mung bean germination, are reported to be products of limited carboxylterminal proteolysis of the major inhibitor MBI-F present in the resting seed (58).

HGI-III, like red kidney bean BBI, exists as a dimer even at lower concentrations (2  $\mu$ g/ml) (25). HGGI-III, derived from HGI-III, exhibits no tendency to associate. The relative mobility on SDS-PAGE indicates that it is a monomer with a molecular mass of 6.5 kDa. A similar behavior is observed in sizeexclusion chromatography (Fig. 4). This distinguishes HGGI-III from HGI-III in its state of association. The amino acid sequence comparison of the two reveals the absence of 12 amino acid residues at the amino terminus and four residues at the carboxyl terminus. The HGI-I, despite lacking the amino terminus, exists as a dimer (18). These observations point to the existence of a structural element, responsible for the self-association of HGI-III, in the deleted tetrapeptide -SHDD. The general features of the Far and Near-UV CD spectra of both HGI-III and HGGI-III are similar<sup>2</sup> indicating they do not differ in their secondary or tertiary structure. The in vitro synthesized BBI and related soybean inhibitor exhibit the phenomenon of self-association (59). The three-dimensional structural data obtained with the BBI from pea seed have clearly shown the last 11 carboxyl-terminal residues of the molecule in the near perfect 2-fold symmetric dimer (5). The residues 68-70 constitute an extended  $\beta$ -strand, which makes no contact with its own subunit. The carboxyl-terminal tail plays a major role in the dimeric association of PsTI-I and PsTI-II (5). Arg<sup>23</sup> and  $Lys^{16}$  (P<sub>1</sub> residue of the first reactive site) of one subunit are involved in the intersubunit contacts with the carboxyl-termi-

<sup>2</sup> P. Kumar, A. G. Appu Rao, and L. R. Gowda, unpublished results.

nal Glu<sup>68</sup> and Glu<sup>69</sup> of the second subunit. The hydrophobic residues that are exposed in the monomeric BBI of black-eyed pea inhibitors are buried in the multimers. The anomalous distribution of hydrophilic and hydrophobic amino acids does not thermodynamically favor their existence as monomers. The light scattering data also support such self-association of BBI from black-eyed pea (15).

SDS-PAGE analysis (Fig. 6) and size-exclusion studies (Fig. 4) of HGI-III, in the presence of 1 mM  $Zn^{2+}$ , further bring out the role of Asp residues at the carboxyl terminus. The carboxyl terminus of HGI-III, -His<sup>74</sup>-Asp<sup>75</sup>-Asp<sup>76</sup>, could ligate  $Zn^{2+}$  in an indirect carboxylate-metal coordination, disrupting the interactions responsible for self-association. The indirect carboxylate-His- $Zn^{2+}$  interaction occurs within several unrelated proteins wherein a His residue bridges both  $Zn^{2+}$  and carboxylate side chain of a nearby Asp (40).

Multiple sequence alignment of legume BBIs reveal that the  $P_1$  residue and the first reactive site in all the BBIs is one of Arg, Lys, or Ala (Fig. 7). This is consistent with the inhibition of trypsin or elastase at the first reactive site. This residue for HGI-III is Lys<sup>24</sup> (19). The Arg<sup>31</sup> of HGI-III, corresponding to Arg<sup>23</sup> of PsTI-IVb, is conserved in only a few BBIs (Fig. 7). In other BBIs, it is replaced by His, Ser, Glu, or Val. These observations indicate that the self-association to form dimers involves intersubunit contact of Lys<sup>24</sup> or Arg<sup>31</sup> in HGI-III.

The role of Lys<sup>24</sup> and Arg<sup>31</sup> of HGI-III in the self-association was probed through a chemical modification approach and further confirmed by homology modeling (Fig. 8). The modification of Arg residues using 1,2-cyclohexanedione did not affect the state of association where as modification of Lys residues using citraconic anhydride resulted in the monomeric form of HGI-III (Fig. 5). These results confirm the involvement of Lys residues in the interaction. The replacement of Arg<sup>23</sup>, which contributes to the dimeric association in PsTI-IVb by Phe in BBI and Ala in A-II, did not affect the state of association. All the BBIs that inhibit elastase at the first reactive site wherein Lys/Arg is replaced by Ala exist as monomers in solution (Table I). The BBIs that inhibit trypsin at the first reactive site (having Lys/Arg) but lack the carboxyl terminal Asp/Glu residues exist as monomers. HGGI-I, -II, and -III, which inhibit trypsin and are lacking in the carboxyl terminal (Fig. 1), exist as



FIG. 8. Interactions at the dimer interface in HGI-III. One monomer is shown in *red* while the other is *blue*. Hydrogen bonds are indicated by *dashed lines*. The salt bridges made by the carboxyl-terminal Asp (Asp<sup>76</sup>) in both subunits with Arg (Arg<sup>31</sup>) from the opposite subunits can be clearly seen. The hydrogen bond between the carboxyl-terminal Asp<sup>76</sup> in both subunits with Lys<sup>24</sup> from the opposite subunit is also visible. The *inset* represents the Ca atoms of the entire dimer by a *ribbon diagram*. The two subunits are labeled A and B, and their amino and carboxyl termini are indicated.



FIG. 9. Thermal stability profiles of HGGI-III and HGI-III at 95 °C. HGGI-III (- $\oplus$ -) and HGI-III (- $\bigcirc$ -).

monomers in solution (24). The BBIs that inhibit trypsin at the first reactive site and possess Asp/Glu residues at the carboxyl terminal exist as dimers in solution (Table I). These observations strengthen the premise that the interaction between Lys and Asp is the essence of self-association of HGI-III. This observation can be further extended to other BBIs whose primary structure and state of association are reported (Table I).

The structure of soybean BBI determined by two-dimensional  ${}^{1}\text{H}$  nuclear magnetic resonance spectroscopy and dynamical simulated annealing (10) also reveals that the antichymotryptic domain at the carboxyl terminus is fully exposed and is presumably the location of the self-association surface of BBI. The second trypsin inhibitory domain of P-II, analogous to the chymotrypsin inhibitory domain, has numerous crystal contacts between protein molecules in the trimers (10) in sup-

port of the notion that the antichymotrypsin domain at the carboxyl terminus is the location of the self-association surface. The dimer model (Fig. 8) clearly discerns the involvement of  $Asp^{76}$ , the carboxyl-terminal residue of HGI-III, in the contacts between the monomers. Of these the hydrogen bond between  $Lys^{24}$  and  $Asp^{76}$  provides the required orientation for the formation of the  $Arg^{31}$ - $Asp^{76}$  salt bridge between the subunits.

The association of BBIs caused by this unique interaction between two monomers must be only of the monomer  $\leftrightarrow$  dimer type with little or no higher forms present as is observed from the model (Fig. 8). The self-association of soybean BBIs characterized by membrane osmometry was found to be of the monomer-dimer type only (4). Considerably high salt concentrations are required to disrupt the HGI-III dimers.<sup>2</sup> NMR analysis indicates that MSTI, a BBI of Snail Medic seeds and a monomer in solution (Table I), can undergo self-association at inhibitor concentrations higher than 2 mm. The residues involved in this mechanism are localized at opposite faces of the molecule, having the highest positive and negative potential. It is to be emphasized that the interactions involved in such self-association (48), which are concentration-dependent are different from the self-association described here. The hydrogen bond observed between Lys<sup>24</sup> (of one subunit) and Asp<sup>76</sup> (of the second subunit) in the dimer (Fig. 8) is the determinant for self-association of HGI-III. In the presence of denaturing agents like 6 M GuHCl, which disrupts such interactions, HGI-III exists as a monomer (18).

The dimer formation may sterically block the chymotrypsinbinding site on one monomer. This, probably accounts for the number of chymotrypsin bound to HGIs determined by activity titration, being less than unity (18). The HGGIs, which are monomers show a 1:1 ratio for both trypsin and chymotrypsin, respectively (24). HGI-III is highly homologous with HGGI-III, but lacks the carboxyl-terminal tetrapeptide -SHDD. The carboxyl-terminal portion of BBIs has been suggested to stabilize conformation of the second reactive site and protect it from enzyme attack. The dimer is more thermostable than HGGI-III (Fig. 9). Most of the exposed hydrophobic residues of the monomeric PsTI-IVb (5) and BTCI (15) are shielded from the solvent by the other subunit in the dimer. This probably accounts for the increased thermal stability of the HGI-III and is indicative of a functional aspect of the dimerization. From these data, we may infer that the degradation of the inhibitor follows the scheme: dimer  $\rightarrow$  monomer  $\rightarrow$  peptides/amino acids  $\rightarrow$  N<sub>2</sub> sink source.

We have demonstrated that using alternate approaches to site-directed mutagenesis, like producing deletion variants of the inhibitor by germination, can also be used to identify the role of individual amino acid residues in the structure and stability of the inhibitor. The results presented in this study clearly support the self-association in HGIs being due to a unique interaction between the  $\epsilon$ -amino group of Lys<sup>24</sup> (P<sub>1</sub> of the first reactive site) of one monomer and the carboxy-side chain of Asp<sup>76</sup>, of the other monomer. Furthermore, this unique interaction also occurs in all the BBIs that self-associate to form dimers. This self-association is vital to the physiological and functional role of BBIs, because it leads to increased thermal stability and greater resistance to enzyme attack by stabilizing the conformation of the second reactive site at the carboxyl-terminal domain.

Acknowledgments—We are grateful to Dr. V Prakash, Director, Central Food Technological Research Institute (CFTRI), Mysore for advice and useful suggestions during the course of this investigation. We also thank Dr. D. Rajagopal Rao, Scientit (Ret.), CFTRI, Mysore, Dr. H. S. Savithri, Department of Biochemistry, and Dr. M. R. N. Murthy, Molecular Biophysics Unit, Indian Institute of Science for their valuable suggestions. We acknowledge the help of V. Sarvanan (Department of Biotechnology facility for protein/peptide sequencing), Indian Institute of Science, Bangalore for technical help in amino-terminal sequence analysis. Our thanks also to P. S. Kulashekhar for help in the preparation of the manuscript.

#### REFERENCES

- 1. Laskowski, M., Jr., and Kato, I. (1980) Annu. Rev. Biochem. 49, 593-626
- 2. Birk, Y. (1985) Int. J. Peptide Protein Res. 25, 113-131
- Ikenaka, T., and Norioka, S. (1986) in *Proteinase inhibitors* (Barret, A. J., and Salvenson, G., eds) pp. 361–374, Elsevier, Amsterdam
- Salvenson, G., eds) pp. 361–374, Elsevier, Amsterdam 4. Harry, J. B., and Steiner, R. F. (1969) *Eur. J. Biochem.* **16**, 174–179 5. Sierra, I., Li de La, Quillien, L., Flecker, P., Gueduen, J., and Brunie, S. (1999)
- J. Mol. Biol. **285**, 1195–1207 6. Kennedy, A. R. (1998) Am. J. Clin. Nutr. **68**, (suppl.) 1406S–1412S
- 7. Harms-Ringdahl, M., Forsberg, J., Fedorcsak, I., and Ehrenberg, L. (1979)
- Biochem. Biophys. Res. Commun. 86, 492–499 8. Chen, P., Rose, J., Love, R., Wei, C. H., and Wang, B. C. (1992) J. Biol. Chem.
- 267, 1990–1994
   Suzuki, A., Yamane, T., Ashida, T., Norioka, S., Hara, S., and Ikenaka, T. (1993) J. Mol. Biol. 234, 722–734
- 10. Werner, M. H., and Wemmer, D. E. (1992) Biochemistry **31**, 999–1010
- Voss, R. H., Ermler, U., Essen, L. O., Wenzl, G., Kim, Y. M., and Flecker, P. (1996) Eur. J. Biochem. 242, 122–131
- Tsunogae, Y., Tanaka, I., Yamane, T., Kikkawa, J., Ashida, T., Ishikawa, C., Watanabe, K., Nakamura, S., and Takahashi, K. (1986) J. Biochem. (Tokyo) 100, 1637–1646
- 13. Lin, G., Bode, W., Huber, R., Chi, C., and Engh, R. A. (1993) Eur. J. Biochem.

212, 549-555

- Koepke, J., Ermler, U., Warkentin, E., Wenzl, G., and Flecker, P. (2000) J. Mol. Biol. 298, 477-491
  - de Freitas, S. M., de Mello, L. V., da Silva, M. C., Vriend, G., Neshich, G., and Ventura, M. (1997) FEBS Lett. 409, 121–127
  - Song, H. K., Kim, Y. S., Yang, J. K., Moon, J., Lee, J. Y., and Suh, S. W. (1999) J. Mol. Biol. 293, 1133–1144
  - Luckett, S., Garcia, R. S., Barker, J. J., Konarev, A. V., Shewry, P. R., Clarke, A. R., and Brady, R. L. (1999) J. Mol. Biol. 290, 525–533
  - Sreerama, Y. N., Das, J. R., Rao, D. R., and Gowda, L. R. (1997) J. Food Biochem. 21, 461–477
  - Prakash, B., Selvaraj, S., Murthy, M. R. N., Sreerama, Y. N., Rao, D. R., and Gowda, L. R. (1996) J. Mol. Evol. 42, 260–569
  - Sreerama, Y. N., and Gowda, L. R. (1997) Biochim. Biophys. Acta 1343, 235–242
  - Ramasarma, P. R., Appu Rao, A. G., and Rao, D. R. (1995) *Biochim. Biophys.* Acta 1248, 35–42
  - Singh, R. R., and Appu Rao, A. G. (2002) Biochim. Biophys. Acta 1597, 280–291
  - Sreerama, Y. N., and Gowda, L. R. (1998) J. Agric. Food Chem. 46, 2596–2600
     Kumar, P., Sreerama, Y. N., and Gowda, L. R. (2002) Phytochemistry 60, 581–588
  - 25. Wu, C., and Whitaker, R. (1990) J. Agric. Food Chem. 38, 1523-1529
  - 26. Bergeron, D., and Neilson, S. S. (1993) J. Agric. Food Chem. 41, 1544-1552
  - Terada, S., Fujimura, S., Kino, S., and Kimato, E. (1994) Biosci. Biotech. Biochem. 58, 371–375
  - Godbole, S. A., Krishna, T. G., and Bhatia, C. R. (1994) J. Sci. Food Agric. 64, 87–93
  - 29. Odani, S., and Ikenaka, T. (1978) J. Biochem. (Tokyo) 83, 747-753
  - Ventura, M. M., Mizuta, K., and Ikemoto, H. (1981) An. Acad. Bras. Cienc. 53, 195–201
  - 31. Laemmli, U. K. (1970) Nature 227, 680–685
  - Aitken, A., Geisow, M. J., Findlay, J. B. C., Holmes, C., and Yarwood, A. (1989) Protein Sequencing a Practical Approach (Geisow, M. J., and Findlay, J. B. C., eds) p. 43, IRL press, Oxford
  - Allen, G. (1989) Laboratory Techniques in Biochemistry and Molecular Biology: Sequencing of Proteins and Peptides, Elsevier, Amsterdam, p. 73
  - 34. Miles, E. W. (1977) Methods Enzymol. 47, 431-442
  - 35. Smith, E. L. (1977) Methods Enzymol. 47, 156-161
  - Dixon, H. B. F., and Perham, R. N. (1968) *Biochem. J.* 109, 312–314
     Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G., and
  - Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673-4680
    38. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Cryst. 26, 283-291
  - 39. Holm, L., and Sander, C. (1996) J. Mol. Biol. 233, 123–138
  - Christianson, D. W., and Alexander, R. S. (1989) J. Am. Chem. Soc. 111, 6412–6419
  - 41. Corpet, F. (1988) Nucleic Acids Res. 16, 10881–10890
  - Deshimaru, M., Hanamoto, R., Kusano, C., Yoshimi, S., and Terada, S. (2002) Biosci. Biotechnol. Biochem. 66, 1897–1903
  - Funk, A., Weder, J. K. P., and Belitz, H. D. (1993) Zlebensm Unters Forsch 196, 343–350
     Tanaka, A. S., Sampaio, M. U., Mentele, R., Aureswald, E. A., and Sampaio,
  - ranaka, A. S., Sampaio, M. U., Mentele, K., Aureswaid, E. A., and Sampaio, C. A. (1997) J. Prot. Chem. 15, 553–560
  - Tanaka, A. S., Sampaio, M. U., Marangoni, S., de Oliveira, B., Novello, J. C., Oliva, M. L., Fink, E., and Sampaio, C. A. (1996) *J. Biol. Chem.* 378, 273–281
  - 46. Gennis, L. S., and Cantor, C. R. (1976) J. Biol. Chem. 251, 747–753
  - Hwang, D. L. R., Lin, K. T. D., Yang, W. K., and Foard, D. E. (1977) Biochim. Biophys. Acta. 495, 239–382
  - Catalano, M., Ragona, L., Molinari, H., Tava, A., and Zetta, L. (2003) Biochemistry 42, 2836–2846
  - 49. Brown, W. E., and Ryan, C. A. (1984) Biochemistry 23, 3418-3422
  - Asao, T., Imai, F., Tsuji, I., Tashoro, M., Iwami, K., and Ibuki, F. (1991) Agric. Biol. Chem. 55, 707-713
  - 51. Wilson, K. A., and Laskowski, M. Sr. (1973) J. Biol. Chem. 248, 756-762
  - 52. Odani, S., and Ikenaka. T. (1977) J. Biochem. 82, 1523-1528
  - Zhang, Y. S., Lo, S. S., Tan, F. L., Chi, C, W., Xu, L. X., and Zhang, A. L. (1982) Sci. Sinica 25, 268–277
  - 54. Joubert, F. J. (1984) Phytochemistry 23, 957-961
  - Joubert, F, J., Kruger, H., Townshend, G. S., and Botes, D. P. (1979) Eur. J. Biochem. 97, 85-91
  - 56. Haynes, R., and Feeney. R. E. (1967) J. Biol. Chem. **242**, 5378–5385
  - Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia C. (1995) J. Mol. Biol. 247, 536–540
  - 58. Wilson, K, A., and Chen, J. C. (1983) Plant Physiol. 71, 341-349
  - Foard, D. E., Gutay, P. A., Ladin, B., Beachy, R. N., and Larkins, B. A. (1982) *Plant Mol. Biol.* 1, 227–243

## Molecular Mechanism of Dimerization of Bowman-Birk Inhibitors: PIVOTAL **ROLE OF ASP76 IN THE DIMERZATION**

Pradeep Kumar, A. G. Appu Rao, Sridhar Hariharaputran, Nagasuma Chandra and Lalitha R. Gowda

J. Biol. Chem. 2004, 279:30425-30432. doi: 10.1074/jbc.M402972200 originally published online April 28, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M402972200

Alerts:

- When this article is citedWhen a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 56 references, 7 of which can be accessed free at http://www.jbc.org/content/279/29/30425.full.html#ref-list-1