J. Agric. Food Chem. **2001**, *49*, 3443−3451

Rapid Gastric Fluid Digestion and Biochemical Characterization of Engineered Proteins Enriched in Essential Amino Acids

Keith R. Roesler and A. Gururaj Rao*

Pioneer Hi-Bred International, Inc., 7300 NW 62nd Avenue, P.O. Box 1004, Johnston, Iowa 50131-1004

The barley high lysine (BHL) proteins are nutritionally enhanced derivatives of barley chymotrypsin inhibitor-2 (CI-2). A compactly folded new CI-2 derivative, BHL9, was engineered with the highest content of threonine, tryptophan, and isoleucine yet achieved in this protein family (15.1, 9.4, and 12.1 wt %, respectively). BHL9 had an unfolding midpoint of 5.5 M guanidinium chloride, significantly greater than values for wild type (3.9 M) or for the previously most stable BHL protein, BHL8 (3.6 M). BHL9 and all other derivatives were digested within 15 s in simulated gastric fluid (SGF), suggesting nutritional availability upon ingestion. Denaturation of the proteins in SGF minus pepsin was revealed by changes in their fluorescence emission spectra and/or far UV circular dichroism spectra. The proteins lack homology to known allergens. Significantly, the BHL8 and BHL9 proteins were stable to proteases at pH 7.5 or 8.0, attesting to their potential for high expression in plants.

Keywords: *Gastric fluid; intestinal fluid; protein digestion; nutritional proteins; protein engineering; protein stability*

INTRODUCTION

The protein of major cereal crops does not contain an optimum amino acid balance for animal or human nutrition. This suboptimal balance limits livestock growth and health (*1*). One approach to improving the nutritional value of crops is to overexpress proteins containing a high content of the limiting amino acids. With this aim in mind, several proteins derived from barley chymotrypsin inhibitor-2 (CI-2) were previously engineered for an increased essential amino acid content (*2*, *3*). The most promising of these engineered proteins, BHL8 (barley high lysine protein 8) was substantially enriched in lysine, methionine, tryptophan, and threonine, retained no inhibitory activity against digestive enzymes of monogastric animals, and contained an engineered disulfide bond that increased stability against proteolysis by trypsin and chymotrypsin. Such in vitro protein stability was a reassuring indication of structural integrity but also raised the question of whether BHL8 would be nutritionally available when eaten. One goal of the present study was therefore to assess the digestibility of the BHL proteins by simulated gastric fluid and simulated intestinal fluid.

Another goal of the present study was to assess the versatility of the CI-2 scaffold to meet diverse nutritional needs. Even a promising protein such as BHL8 is not optimal for solving amino acid requirements in all cases. As compared to poultry, for example, swine have a higher requirement for tryptophan and a lower requirement for the sulfur amino acids (*1*). Furthermore, the relative amino acid requirements vary with the age of the animal. In addition, the relative content of essential amino acids varies between conventional maize and high oil maize, and between maize and the other cereals. For example, maize protein has only approximately half the tryptophan content of protein from the other cereals (*4*). Finally, the relative price and commercial availability of free amino acids for use as feed supplements can influence the amino acid goals. It is apparent, therefore, that a single engineered protein would not be optimal for meeting amino acid goals in all situations. A flexible protein platform that could be tailored for a specific amino acid content, depending on the intended use, would be helpful. To this end, we have explored the flexibility of the CI-2 protein as a platform for nutritional enhancement of maize and other crops.

We here report the design of another disulfide stabilized CI-2 derivative, BHL9, that has greater in vitro protein stability and a higher content of threonine, tryptophan, and isoleucine, in comparison with BHL8. In addition, we demonstrate that all members of the BHL protein family are readily digested by enzymes of the gastrointestinal tract of monogastric animals.

MATERIALS AND METHODS

BHL9 Expression and Purification. A synthetic gene encoding the BHL9 protein was prepared by Midland Certified Reagent Company (Midland, TX) and ligated into the *Nco*I and *Hin*dIII restriction sites of expression vector pET 28 (Novagen, Madison, WI). BHL9 was expressed in *Escherichia coli* BL21 DE3 in 2× YT media. Cultures were grown at 37 °C to an $OD_{600\ nm}$ of 0.6 to 0.8, then induced with 1 mM IPTG and transferred to 16 °C for growth overnight. The protein was purified by cation exchange and gel permeation chromatography as reported previously (*3*), except that the gel permeation chromatography was done in 50 mM Tris-HCl, 150 mM NaCl, pH 8.6. N-terminal amino acid sequencing and MALDI mass spectrometry analyses of purified BHL9 were performed by the Iowa State University Protein Facility (Ames, IA).

SDS−PAGE and Immunoblot Analysis. SDS−PAGE was performed with precast Tris-Tricine 16.5% gels from BioRad Laboratories (Hercules, CA) as per manufacturer's instructions. Proteins were visualized by Coomassie brilliant

* To whom correspondence should be addressed (phone: (515) 270-3576; fax: (515) 253-2149; e-mail: raog@phibred.com).

10.1021/jf010209b CCC: $20.00 © 2001 American Chemical Society
Published on Web 06/19/2001

blue. For the immunoblots of Figures 6 and 9, protein from SDS−PAGE was blotted to Immobilin-P PVDF membrane from Millipore (Bedford, MA) with a Trans-Blot SD semidry transfer cell (BioRad Laboratories, Hercules, CA). Blots were blocked in 3% bovine serum albumin 15 min, then probed with a 1:1000 dilution of rabbit anti-sera against either wild-type CI-2 (Figure 9 blot and wild-type blot of Figure 6) or against a 1:1 mixture of BHL6 (*2*) and BHL8 (BHL8 blot of Figure 6). Proteins were visualized by probing with a goat anti-rabbit alkaline phosphatase conjugate according to the method of the manufacturer (Promega Biotec, Madison, WI).

BHL Protein Quantitation. A molar extinction coefficient at 280 nm of 23 000 M^{-1} cm^{-1} was used for BHL9 quantitation, and previously determined coefficients were used for the other BHL and wild-type proteins (*2*, *3*).

Protease Inhibition Assays. Apparent K_i values for BHL9 were determined with the proteases and method described previously (*3*).

Trypsin and Chymotrypsin Digests. Bovine pancreatic trypsin and chymotrypsin from Sigma Chemical Company (St. Louis, MO) were quantitated by absorbance at 280 nm using an extinction coefficient (1% protein, 1-cm path length) of 16.0 for trypsin and 20.4 for chymotrypsin. Digests were done at 37 °C for 4 h in 100 mM Tris-HCl, 50 mM NaCl, 2 mM $CaCl_2$, pH 8.0. Three micrograms of BHL8 or BHL9 were incubated with or without 0.3 μg of protease in a 15 μL volume. Reactions were stopped by adding an equal volume of BioRad 2× Tris-Tricine SDS sample buffer containing 200 mM dithiothreitol and 6 mM PMSF, followed by boiling 5 min and SDS−PAGE.

Simulated Gastric Fluid Digests. SGF contained 3.2 mg/mL porcine pepsin (Sigma Chemical Company, St. Louis, MO) in 34 mM NaCl, 0.7% HCl, pH 1.2 (*5*). Incubations were done at 37 °C with 15 μL of SGF per 3 μg of target protein. This gave a ratio of 16:1 (wt/wt) pepsin/target protein, which is similar to the 18.8:1 ratio used previously in a study of allergenic and nonallergenic proteins (*6*). At the desired times, 15 μL aliquots of the incubation mix were transferred to a stop solution composed of 15 μL of BioRad 2× Tris-Tricine SDS sample buffer containing 200 mM dithiothreitol, plus 1 μL of 1.5 M Tris-HCl, pH 8.8. For the 0 time controls of Figure 5, SGF and substrate protein were added directly to the stop solution without prior incubation. Digestion was assessed by SDS−PAGE and Coomassie brilliant blue staining. To preincubate BHL8 or WT CI-2 with a maize protein matrix (Figure 6 experiment), whole kernel flour from opaque-2 maize was extracted 4 h at 25 °C in a 10:1 vol/wt ratio of 50 mM Tris-HCl, 100 mM KCl, 5% glycerol, pH 8.0. Maize protein was quantitated by the method of Bradford (*7*) using the reagent from BioRad Laboratories (Hercules, CA) with bovine gamma globulin as standard. A 20:1 wt/wt ratio of maize protein/BHL8 or WT CI-2 was incubated 30 min at 37 °C before proceeding with the SGF digests.

Determination of Initial BHL8 Peptic Cleavage Sites. BHL8 (60 μg) was digested 15 min at 37 °C in 34 mM NaCl, 0.7% HCl, pH 1.2, with a 1:50 (wt/wt) ratio of pepsin/BHL8. The peptic fragments were purified by reversed phase chromatography with a PepRPC HR 5/5 column (Amersham-Pharmacia; Piscataway, NJ). Elution was performed in 0.1% TFA with a 30 mL, 10−45% acetonitrile gradient. N-terminal sequencing and MALDI mass spectrometry were performed by the Iowa State University Protein Facility.

Simulated Intestinal Fluid Digests. SIF is defined as 50 mM potassium phosphate, pH 7.5, containing 10 mg/mL pancreatin (*5*). However, the porcine pancreatin used in the present study (Sigma Chemical Company, St. Louis, MO catalog no. P-7545) was 8× United States Pharmacopeia specifications, and therefore, 1.25 mg/mL, rather than 10 mg/mL, was used. Incubations were done at 37 °C, and at the desired times, aliquots containing 3 μg of BHL or wild-type protein in 15 μL of SIF were transferred to a stop solution composed of 16 μL of BioRad 2× Tris-Tricine SDS sample buffer containing 200 mM dithiothreitol and 6 mM PMSF, and immediately boiled one minute. The 0 time controls were done by adding SIF and substrate proteins directly to stop solution, without prior incubation.

Digestion of BHL8 Peptic Fragment by Trypsin or Simulated Intestinal Fluid. To generate the BHL8 peptic fragment, BHL8 (20 μg) was digested with 0.4 μg of porcine pepsin in 40 μL of 34 mM NaCl, 0.7% HCl, pH 1.2, at 37 °C for 15 min. The reaction was stopped by adding 40 μL of 1 M Tris-HCl, pH 8.0. Twenty microliters of the 80 μL total volume was transferred to SDS sample buffer and represented lane 2 of Figure 7. Another 20 μL of the peptic fragment was further digested by addition of 0.5 μg of trypsin and incubation at 37 °C 30 s, and the reaction was stopped by addition of an equal volume of BioRad 2× Tris-Tricine SDS sample buffer containing 200 mM dithiothreitol and 5 mM PMSF. This digest represented lane 3 of Figure 7. An additional 20 μL aliquot of the peptic fragment was further digested by incubation with SIF at 25 °C 30 s, and the reaction was stopped as done with trypsin. This digest represented lane 4 of Figure 7.

Purification and Quantitation of CI-2 Homologues from Mature Maize Kernels. Twenty grams of opaque-2 maize whole kernel flour was extracted in 100 mL of 70% ethanol 18 h at 25 °C. Centrifugation was done at 17000*g* for 15 min, and the supernatant was saved. The pellet was back extracted with an additional 70 mL of 70% ethanol, and centrifugation was repeated. The two supernatants were combined and poured through one layer of Miracloth (Calbiochem-Novabiochem, La Jolla, CA). Ethanol was then added to 90%, and stirring was done at 25 °C, 40 min. Centrifugation was done at 17000*g* 15 min. The pellet was resuspended in 20 mL of 50 mM Tris-HCl, pH 8.0, and recentrifuged to remove insoluble material. The supernatant, representing the 70−90% ethanol fraction, was applied to a protease-agarose (Sigma Chemical Company, St. Louis, Mo.) column containing 2 units of Protease VIII A from *Bacillus licheniformis*. The column was washed with 10 column volumes of 50 mM Tris-HCl, pH 8.0, then eluted with 5 column volumes of 100 mM glycine, pH 2.0. Neutralization was done with an equal volume of 1 M Tris-HCl, pH 8.0. Buffer was changed to 10 mM sodium phosphate, pH 7.0, with a disposable PD-10 column from Amersham-Pharmacia, and the protein was concentrated by ultrafiltration with Centricon-3 concentrators (Millipore; Bedford, MA). A total of 164 μg of affinity purified CI-2 homologues resulted, as determined by absorbance at 205 nm using an extinction coefficient (1 mg/mL) of 27 + 120 (A_{280}/A_{205}) (*8*). This quantitation method is known to be relatively independent of amino acid composition of the protein. Assuming the 20 g of flour was 10% protein, and assuming a 10% recovery during extraction, purification and concentration, the CI-2 homologues collectively comprised about 0.08% of total kernel protein.

Circular Dichroism and Fluorescence. Far UV CD and fluorescence emission spectra were determined as described previously (*2*) except that protein concentrations of 5 μM, rather than 2 μM, were used for fluorescence analyses. For the samples in simulated gastric fluid minus pepsin, CD and fluorescence analyses were begun 5 min and 30 s, respectively, after addition of the protein to the pH 1.2 buffer.

Homology Search of Allergen Database. A TFASTA search (*9*) of the Genetics Computer Group, Inc. (GCG, Madison, WI) Allergens database was made using a gap creation penalty of 16 and a gap extension penalty of 2.

RESULTS

Design of BHL9. The BHL9 composition and amino acid sequence are presented in Table 1 and Figure 1, in comparison with WT CI-2 and other BHL proteins. The intent with BHL9 was to make changes in the BHL8 composition favorable for swine nutrition, without disrupting protein folding and stability. Thus, the two cysteine substitutions at positions 22 and 82 that formed a stabilizing disulfide bond in BHL8 (*2*) were retained in BHL9. Threonine, tryptophan, and isoleucine were increased in BHL9, at the expense of some methionine and lysine. The methionine reduction was considered acceptable because this amino acid is far less limiting

Table 1. Amino Acid Composition of Wild-Type and Engineered CI-2

AA	WT CI-2 #	WT CI-2 mol%	WT CI-2 wt%	BHL8 #	BHL8 mol%	BHL8 wt%	BHL9 #	BHL9 mol%	BHL9 wt%	average proteins[a] mol%	maize endosperm proteins[b] wt%
Ala	3	4.6	3.1	4	6.0	4.0	4	6.0	4.1	8.3	8.1
Arg	4	6.1	8.0	2	3.0	3.9	2	3.0	4.0	5.7	3.8
Asn	2	3.0	3.0	0	0	0	0	0	0	4.4	6.2 (Asn+Asp)
Asp	4	6.1	6.1	4	6.0	6.0	4	6.0	6.1	5.3	
Cys	0	0	0	2	3.0	2.7	2	3.0	2.8	1.7	1.8
Gln	3	4.6	5.0	0	0	0	0	0	0	4.0	21.3 (Gln+Glu)
Glu	6	9.1	10.1	4	6.0	6.6	3	4.5	5.1	6.2	
Gly	3	4.6	2.6	4	6.0	3.4	4	6.0	3.5	7.2	3.2
His	0	0	0	0	0	0	0	0	0	2.2	2.8
Ile	6	9.1	9.0	3	4.5	4.4	8	11.9	12.1	5.2	3.8
Leu	6	9.1	9.0	3	4.5	4.4	2	3.0	3.0	9.0	14.3
Lys	6	9.1	10.1	16	23.9	26.3	10	14.9	16.8	5.7	2.0
Met	2	3.0	3.4	6	9.0	10.1	2	3.0	3.4	2.4	2.8
Phe	1	1.5	1.9	0	0	0	0	0	0	3.9	5.3
Pro	4	6.1	5.3	4	6.0	5.2	4	6.0	5.3	5.1	9.7
Ser	1	1.5	1.2	0	0	0	0	0	0	6.9	5.2
Thr	3	4.6	4.1	5	7.5	6.7	11	16.4	15.1	5.8	3.5
Trp	1	1.5	2.4	3	4.5	6.9	4	6.0	9.4	1.3	0.5
Tyr	1	1.5	2.1	0	0	0	0	0	0	3.2	5.3
Val	10	15.2	13.5	7	10.5	9.2	7	10.5	9.4	6.6	4.7
Total	66	100.3	99.9	67	100.4	99.8	67	100.2	100.1	100.1	104.3

[a] Average of 1021 unrelated proteins (*28*). [b] Glover and Mertz (*21*).

```
                                  10        20        30        40        50        60        70        80
                                   *         *         *         *         *         *         *         *
CI-2(native)     SSVEKKPEGVNTGAGDRHNLKTEWPELVGKSVEEAKKVILQDKPEAQIIVLPVGTIVTMEYRIDRVRLFYDKLDNIAQVPRVG
CI-2(truncated)                   MNLKTEWPELVGKSVEEAKKVILQDKPEAQIIVLPVGTIVTMEYRIDRVRLFYDKLDNIAQVPRVG
BHL9                             MATTKCTWPELVGKTVEWAKKIIMKDKPTATIIVIPVGTIVTGEWKIDRVRLWVDKTDTIAKTPKCG
BHL8                             MAKMKCTWPELVGKTVEKAKKMIMKDKPEAKIMVLPVGTKVTGEWKMDRVRLWVDKKDKIAKTPKCG

BHL6                             MAKMKTTWPELVGKTVEKAKKMIMKDKPEAKIMVLPVGTKVTGEWKMDRVRLWVDKKDKIAKTPKVG
BHL5                             MAKMKTTWPELVGKTVEKAKKMIMKDKPEAKIMVLPVGTKVTGEWKMDRVKLWVDKKDKIAKTPKVG
BHL4                              MKLKTEWPELVGKSVEKAKKVILKDKPEAQIIVLPVGTKVTGEYKIDRVKLFVDKKDNIAQVPRVG
BHL3N            MKSVEKKPKGVKTGAGDKHKLKTEWPELVGKSVEKAKKVILKDKPEAQIIVLPVGTKVGKHYKIDKVKLFVDKKDNIAQVPRVG
BHL3                              MKLKTEWPELVGKSVEKAKKVILKDKPEAQIIVLPVGTKVGKHYKIDKVKLFVDKKDNIAQVPRVG
BHL2                              MKLKTEWPELVGKSVEKAKKVILKDKPEAQIIVLPVGTKVAKAYKIDKVKLFVDKKDNIAQVPRVG
BHL1                              MKLKTEWPELVGKSVEKAKKVILKDKPEAQIIVLPVGTKVTKEYKIDRVKLFVDKKDNIAQVPRVG

Maize 1                 MSSTECGGGGGGAKTSWPEVVGLSVEDAKKVILKDKPDADIVVLPVGSVVTADYRPNRVRIFVDIV---AQTPHIG
Maize 2                 MSSVVLGATGRENKTSWPEVVGMSIKEAREIILKDMPNANIQVLPVGSLVTQDFRPDRVRIFVDIV---AQTPTVG
Maize 3              MSSVVVDAAATSSSEKKTSWPEVVGMSIKEATETILKDMPNAYIQVLPVGSPVTLDLRPDRVRIFVDTV---AMTPTVG
```

Figure 1. Amino acid sequences of engineered and wild-type barley CI-2 and homologous sequences from maize endosperm. Mutated residues are in bold. Arrows indicate the initial peptic cleavage sites of BHL8. Underlined residues were determined by N-terminal sequencing of BHL8 peptic fragments. Dashed lines represent the disulfide bond between Cys22 and Cys82 in BHL8 and BHL9. Native CI-2 sequence is from Williamson et al. (*27*).

for swine than for poultry. Lysine is the first limiting amino acid for swine fed a typical corn/soybean diet (*1*), but lysine is much less expensive as a feed supplement than threonine, tryptophan, or isoleucine. Therefore, lysine substitutions at several positions in BHL8 were resubstituted with threonine, tryptophan, or isoleucine in BHL9. In choosing the BHL9 substitutions, careful consideration was given to an alignment of homologous sequences of CI-2 from diverse plant species (*2*). For example, a tryptophan substitution was made at position 34 of BHL9, even though this is not a conservative change in comparison with BHL8 or WT CI-2, because several homologous sequences had aromatic residues at this position. Similarly, numerous homologues had isoleucine at position 51, and threonine (or serine) at position 45, suggesting that these substitutions in BHL9 were acceptable. The lysine, threonine, tryptophan, sulfur amino acid, and isoleucine contents of the BHL proteins are presented in Table 1. These values are much higher than those of proteins in general, or of maize endosperm protein in particular.

Biochemical Characterization of BHL9. BHL9 was initially expressed in *E. coli* at 37 °C and purified with the same method as used previously for the other BHL proteins. However, much lower yields were consistently observed with BHL9 than with any of the other proteins (<50 μg of BHL9 versus several milligrams of the other proteins per liter of culture). Growth overnight at 16 °C after induction increased BHL9 expression severalfold, but yields were still much lower than for the other BHL proteins at 37 °C. Amino terminal sequencing of purified BHL9 gave the sequence "Ala–Lys–Met–Lys", revealing that the start methionine was not retained in BHL9 (Figure 1), in agreement with previous observations for some of the other BHL proteins.

Protease inhibitory activities of BHL9 were determined with pancreatic chymotrypsin, trypsin, and elastase, and also with subtilisin Carlsberg from *B. licheniformis*. No inhibitory activities against chymotrypsin, trypsin, or elastase were observed, consistent with previous observations for BHL8 and the closely

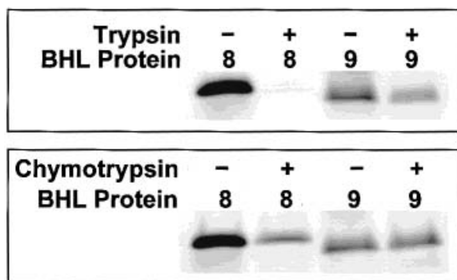

Figure 2. Relative resistance of BHL9 and BHL8 to digestion by trypsin and chymotrypsin. Incubations were done at 37 °C for 4 h without protease or with 1:10 (wt:wt) protease/substrate.

related proteins BHL5 and BHL6 (*2*). However, BHL9 was an effective inhibitor of subtilisin, with an apparent K_i value of 11.6 ± 2.1 nM. This is similar to the WT value but is very different from the BHL8 value of 415 ± 3 nM (*2*). Thus, one or more of the BHL9 amino acid substitutions resulted in a significant restoration of inhibitory activity against this bacterial protease. The WT isoleucines that were restored in BHL9 at positions 56 and 63 are both located in the reactive site loop region and perhaps contributed to this restoration of inhibitory activity.

The BHL9 thermodynamic stability was determined by equilibrium unfolding experiments with guanidinium chloride, as reported previously (*2*). An unfolding midpoint of 5.5 M GdmCl was determined for BHL9, significantly greater than the previously determined values of 3.9 M for wild type CI-2 and 3.6 M for BHL8, the most thermodynamically stable BHL protein previously characterized.

The structural integrity of BHL9 was also examined by incubation with the proteases trypsin and chymotrypsin (Figure 2) and compared with the most proteolytically stable BHL protein identified previously, BHL8. Despite having many potential tryptic cleavage sites (Lys and Arg), BHL9 was relatively resistant to digestion by trypsin, suggesting that BHL9 was compactly folded. In contrast, no BHL8 protein survived the trypsin incubation intact, but this may be due in part to the higher content of lysine (and therefore of potential tryptic cleavage sites) of BHL8, rather than reflecting a difference in compactness of folding. Stability against digestion by chymotrypsin, which cleaves at aromatic residues, was also examined (Figure 2). Partial digestion of BHL8, but not BHL9, was evident. Thus, BHL9 appeared to be more resistant to digestion by chymotrypsin, despite having one additional potential chymotryptic cleavage site at position 34 that was not present in BHL8. The four tryptophans in BHL9 are located in a β-sheet (Trp 24 and Trp 69), an α-helix (Trp 34) and the reactive site loop (Trp 61). The fact that these residues were relatively inaccessible to chymotrypsin suggested that BHL9 was compactly folded.

Changes in Protein Conformation in Simulated Gastric Fluid Minus Pepsin. Wild-type CI-2, BHL9, BHL8, and BHL6 (identical to BHL8 except for the disulfide bond) were analyzed by far UV circular dichroism spectroscopy at pH 7 and also at pH 1.2 in simulated gastric fluid minus pepsin (Figure 3). At pH 7, the BHL9 spectrum was very similar to that of its parent protein BHL8, suggesting similar secondary structures. Both showed a minimum at 207 nm, with a slightly less negative ellipticity for BHL9. Major changes in the BHL6 and BHL8 spectra were observed at pH 1.2 as compared with pH 7, suggesting substantial changes in the secondary structure of the two proteins at low pH. In contrast, only minor differences were observed for the wild type and BHL9 spectra at these two pH values, indicating greater retention of secondary structure in these proteins at low pH. The BHL9 spectrum even appeared to retain the minimum at 234 nm at pH 1.2. This minimum was previously attributed to tryptophan 24 of wild-type CI-2 (*10*).

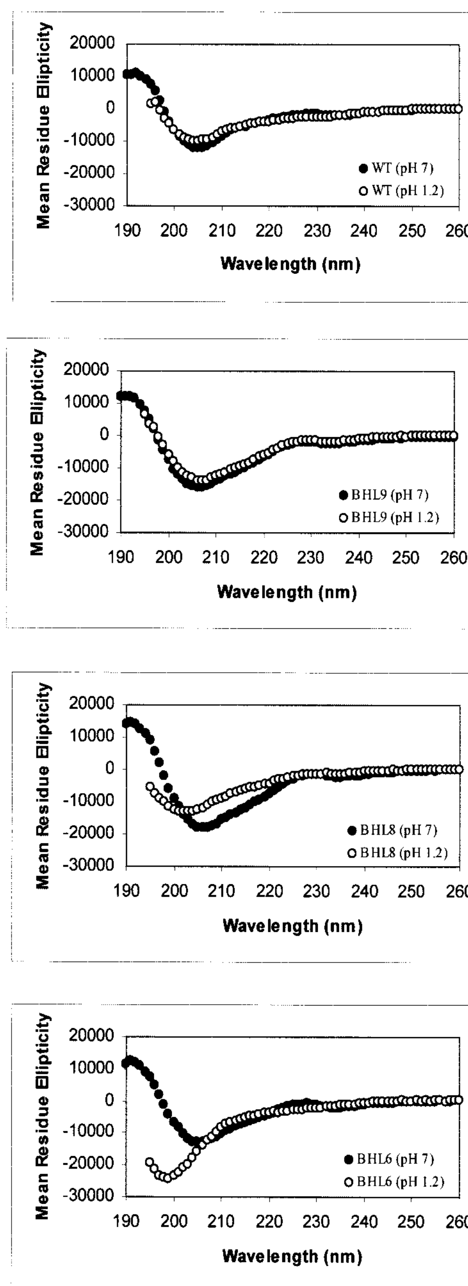

Figure 3. Far UV circular dichroism spectra of wild-type and engineered CI-2. Analyses were done with a protein concentration of 20 μM in either 10 mM sodium phosphate, pH 7, or in simulated gastric fluid minus pepsin (pH 1.2).

The fluorescence emission spectra of the proteins were also examined at neutral pH and in simulated gastric fluid minus pepsin (Figure 4). Immediate changes in fluorescence intensity were observed for all of the proteins upon addition to the SGF buffer, with an increase in intensity noted for wild type CI-2 and

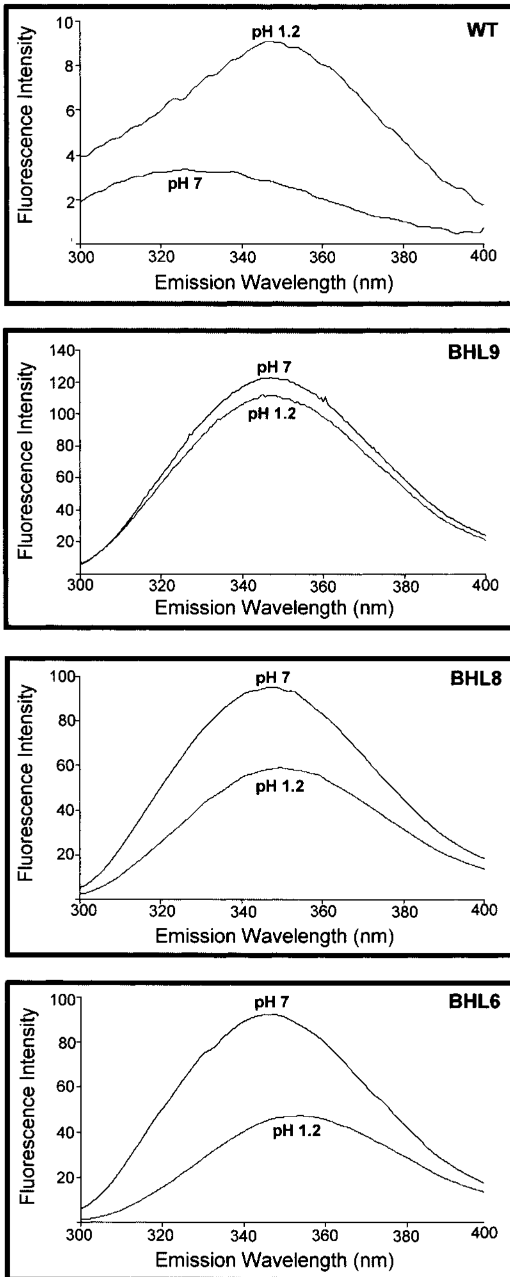

Figure 4. Fluorescence emission spectra of wild-type and engineered CI-2. Excitation wavelength was 280 nm. Analyses were done with a protein concentration of 5 μM in either 10 mM sodium phosphate, pH 7, or in simulated gastric fluid minus pepsin (pH 1.2).

decreases for the other three proteins. Fluorescence intensities were lowest, as expected, in the wild-type protein, which contained only a single tryptophan. Red shifts in the peak emission wavelength at low pH were also observed with all of the proteins except for BHL9, suggesting greater exposure of tryptophan(s) at low pH. Taken together, the results of Figures 3 and 4 suggested that all of the proteins underwent conformational changes at low pH, with the least change occurring in the BHL9 protein.

Digestion by Simulated Gastric Fluid. Digestion of the engineered and wild-type proteins by simulated gastric fluid (SGF) was examined (Figure 5). All of the engineered and wild type proteins were digested by SGF within 15 s, including BHL9, the protein showing the

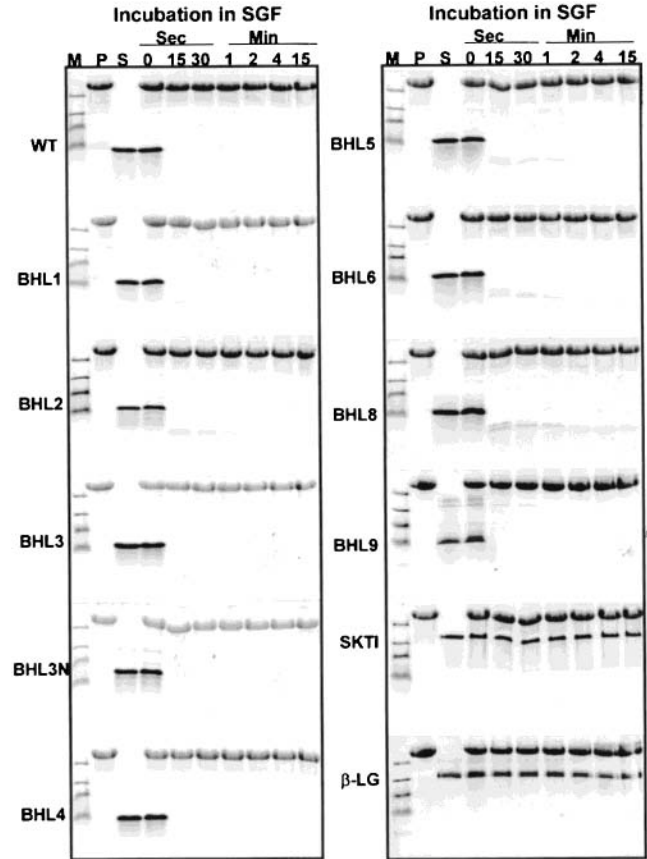

Figure 5. Digestion of wild-type and engineered CI-2 by simulated gastric fluid (SGF). SDS–PAGE and Coomassie brilliant blue staining were used to assess digestion. Lane M, molecular mass markers of 26.6, 17.0, 14.4, and 6.5 kDa. A 3.5 kDa marker is also barely visible in some gels. Lane P, pepsin (no substrate protein); lane S, substrate protein (no pepsin); other lanes, substrate protein incubated at 37 °C in SGF for the indicated times. SKTI = soybean Kunitz trypsin inhibitor. β-LG = beta subunit of lactoglobulin from bovine milk.

least change in CD or fluorescence emission spectra at low pH. In contrast, the soybean Kunitz trypsin inhibitor and the beta subunit of bovine milk lactoglobulin were completely stable in SGF, in agreement with previous results (*6*). For some of the engineered proteins, the peptic fragments were large enough to detect by SDS–PAGE. For example, two peptic fragments of BHL5, BHL6, and BHL8 were visible at 15 s. Only the smaller of the two was detected at 1 min, and this fragment gradually diminished during the time course, but was still barely visible in BHL8 at 15 min. For other proteins such as BHL4, BHL9, and wild-type CI-2, the peptic fragments were too small to be detected by this gel system, even at 15 s.

It is possible, though unlikely, that a maize protein could bind to the engineered proteins and stabilize them against pepsin digestion. For example, perhaps a very acidic maize protein could bind to the basic BHL proteins and directly block peptic cleavage sites or indirectly block them by preventing unfolding of the BHL proteins. This hypothesis was tested by incubation of BHL8 or wild-type CI-2 with a maize protein extract. However, the two proteins were still digested by SGF within 15 s (Figure 6).

The susceptibility of the most persistent BHL8 peptic fragment to further digestion by trypsin or simulated

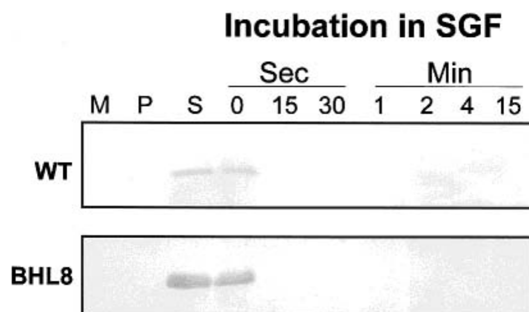

Figure 6. Digestion of wild-type CI-2 and BHL8 by simulated gastric fluid (SGF), following incubation in a corn protein matrix. Pure BHL8 or CI-2 were incubated 30 min in a 20-fold excess of corn protein. SGF digests were then done as in Figure 5, except that immunoblot analysis of gels was done, rather than Coomassie brilliant blue staining.

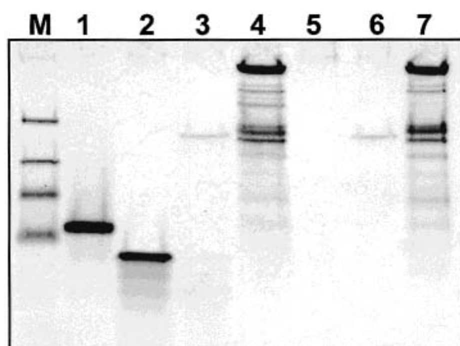

Figure 7. Digestion of BHL8 peptic fragment by trypsin and by simulated intestinal fluid (SIF). SDS–PAGE and Coomassie brilliant blue staining were used to assess digestion. Lane M, molecular mass markers of 26.6, 17.0, 14.4, and 6.5 kDa; lane 1, BHL8 (5 μg); lane 2, BHL8 peptic fragment resulting from digestion of 5 μg of BHL8; lane 3, BHL8 peptic fragment (same quantity as in lane 2) incubated 30 s with trypsin; lane 4, BHL8 peptic fragment (same quantity as in lane 2) incubated 30 s with SIF; lane 5, pepsin only; lane 6, trypsin only; lane 7, SIF only.

intestinal fluid (SIF) was also examined (Figure 7). The BHL8 peptic fragment was digested within 30 s by either trypsin or SIF. Thus, the BHL8 peptic fragment was much more easily digested by trypsin than was the intact BHL8 protein, which was still detectable in 1:10 trypsin after 2 h (*2*).

To determine the initial BHL8 peptic cleavage sites, digestion with a dilute pepsin concentration was performed, and the resulting BHL8 peptic fragments were purified by reversed phase chromatography. Three protein peaks were purified and subjected to N-terminal sequencing and MALDI analysis (Figure 1). One peak had a sequence of "Val--Gly--Lys--Thr" and a mass of 4.7 kDa. Another peak had a double sequence of "Ala, Trp—Lys, Val—Met, Asp—Lys" and a mass of 2.9 kDa, while the third peak represented residual undigested BHL8. The sequences and masses of the peptic fragments were identical to those expected for pepsin cleavage of BHL8 after Leu-27 and Leu-68, and were consistent with the protein containing a disulfide bond, and not retaining the start Met. An examination of the 3-D structure of the folded protein suggested that Leu-27 and Leu-68 are unlikely to be accessible to pepsin, as the side chains of these residues are buried ($\approx$ 0% solvent accessibility). The extremely low pH of SGF apparently denatured the protein, exposing these peptic cleavage sites, consistent with the changes in the BHL8

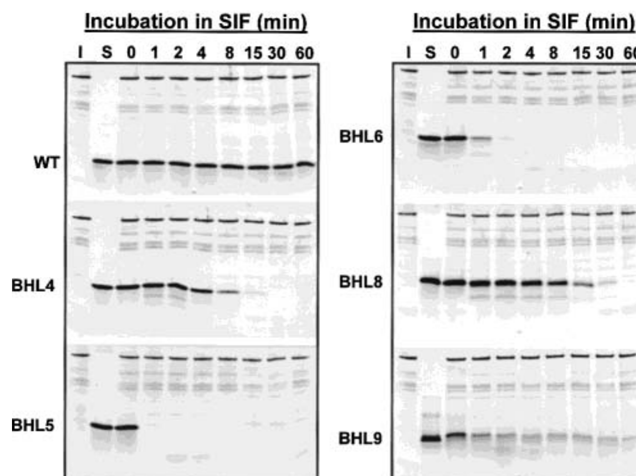

Figure 8. Digestion of wild-type and engineered CI-2 by simulated intestinal fluid (SIF). SDS–PAGE and Coomassie brilliant blue staining were used to assess digestion. Lane I, SIF only; lane S, substrate protein only; other lanes, substrate protein incubated at 37 °C in SIF for the indicated times.

Table 2. Comparison of Thermodynamic and Proteolytic Stability of Wild-Type and Engineered CI-2

protein	unfolding midpoint [GdmCl] (M)[a]	survival time in SIF (min)	survival time in SGF (s)
WT CI-2	3.9	60 (no digestion evident)	<15
BHL9	5.5	60 (some digestion evident)	<15
BHL8	3.6	30	<15
BHL4	2.6	15	<15
BHL6	1.8	2	<15
BHL5	1.3	<1	<15

[a] Value for BHL4 was from Roesler and Rao (*3*). Values for other proteins (except BHL9) were from Roesler and Rao (*2*).

CD spectra and fluorescence emission spectra at low pH noted previously (Figures 3 and 4).

Digestion by Simulated Intestinal Fluid. On the basis of the results of Figure 5, it seemed unlikely that the BHL proteins would survive passage through the stomach to reach the intestine intact. Nevertheless, it was interesting to look at the digestibility of wild type CI-2 and several of the intact BHL proteins by simulated intestinal fluid (SIF) (Figure 8). Wild-type CI-2 was the most stable protein in SIF, with no digestion evident after even 60 min in this potent protease cocktail. BHL9 was the most stable of the engineered proteins, with some intact BHL9 detected at 60 min. Substantial digestion of BHL9 was also evident at 60 min, however, unlike for wild type CI-2. BHL8 and BHL4 were also relatively resistant to digestion by SIF, with intact protein evident at 30 and 15 min, respectively. In contrast, BHL6 and BHL5 were much less resistant to digestion, surviving only 2 min or <1 min, respectively. The thermodynamic stability of the proteins correlated well with their stability in SIF at pH 7.5, but not with their stability in SGF at pH 1.2 (Table 2). The only exception to the correlation was wild type CI-2, which was thermodynamically less stable than BHL9, yet more stable in SIF. This result may be explained by the fact that wild-type CI-2 is a potent inhibitor of chymotrypsin and elastase (*11*), both components of SIF, but BHL9 is not an inhibitor of these proteases.

Assessment of Allergenicity Risk. According to the FAO/WHO report on the evaluation of allergenicity (Evaluation of Allergenicity of Genetically Modified Foods, http://www.fao.org/WAICENT/FAOINFO/ECONOMICS/ESN/gm/biotec-e.htm), the methods include (i) in vitro digestibility in SGF, (ii) sequence homology searches, and (iii) specific serum screening.

In the present paper, we have used the first two methods in the evaluation of BHL proteins. The SGF digests presented in Figure 5 were helpful not only for predicting nutritional availability but also for predicting the risk of food allergenicity. It has been suggested that food allergens tend to be stable in SGF, while nonallergens tend to be quickly digested (*6*). The BHL proteins behaved like nonallergens, being digested within 15 s, while the Kunitz trypsin inhibitor and β-lactoglobulin, which are both known food allergens (*6*), were stable.

Another means of predicting the risk of food allergenicity is to search the protein of interest for sequence identity with known allergens. A protein containing six consecutive amino acids that are identical to the sequence of a known allergen or having 35% identity in an 80-amino acid stretch is considered to be at risk of being an allergen. Therefore, a search of the GCG allergen database was made with the BHL8 and BHL9 amino acid sequences. No matches of greater than four consecutive amino acids were identified nor were overall sequence identities of >35% found. The serum of individuals allergic to barley would be tested for binding of IgE to CI-2 and derivatives only if successful in planta expression of the proteins is achieved.

Overexpressing the engineered proteins in maize would also appear less likely to introduce a new allergen if similar, naturally occurring proteins were already present in maize kernels. To determine the presence of CI-2 homologues in maize, a search with the wild type CI-2 sequence was made of the Pioneer Hi-Bred maize embryo, endosperm, and whole kernel EST databases. EST sequences encoding three CI-2 homologues from maize endosperm were identified (Figure 1). Homologue 1 was also previously determined to be expressed in maize germinating embryos (*12*) and leaves (*13*). The N-terminal regions of the homologues were nonconserved and appeared unstructured, as determined previously with the barley CI-2 N-terminal region (*14*). Over the 65-amino acid region from positions 19 to 83, homologue 1 had amino acid sequence identities of 62, 52, and 51% with WT CI-2, BHL8, and BHL9, respectively. The corresponding values for homologue 2 were 54, 46, and 46%, and the values for homologue 3 were 52, 46, and 45%. To approximately quantitate the levels of the CI-2 homologues in mature opaque-2 maize kernels, a purification method was developed that included ethanol fractionation and affinity chromatography with a protease agarose column (Figure 9). The polypeptide(s) represented by the most prominent band in lane 1 of Figure 9 ran near the 7.5 kDa molecular mass marker, bound to the protease, and cross-reacted with polyclonal antibodies against barley CI-2. These observations were all consistent with expectations for CI-2 homologues. Using assumptions described in Materials and Methods, we estimated that the CI-2 homologues collectively comprised between 0.01 and 0.1% of the total protein in mature maize kernels.

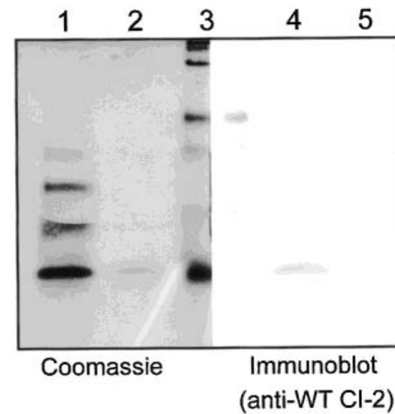

Figure 9. Purification of CI-2 homologues from mature maize kernels. Lanes 1 and 4, protein purified by 70 to 90% ethanol fractionation; lanes 2 and 5, protein affinity purified with a protease-agarose column. Lane 3 (split for both staining and blotting), prestained molecular mass markers of 7.5, 18.4, 32.5, 45.7, 78, 132, and 216 kDa.

DISCUSSION

Nutritional Availability of Engineered Proteins. Our studies suggest that the BHL proteins will be nutritionally available when eaten by humans or monogastric animals. Both the engineered and wild-type proteins were digested by SGF within 15 s, consistent with a previous observation that the CI-2 inhibitory activity against chymotrypsin was eliminated by a 5 min incubation with pepsin at pH 2.0 (*15*). Considering that there is a lag period of 10 to 70 min following ingestion before any solid food is emptied from the human stomach (*16*), it seems unlikely that the BHL proteins, or wild-type CI-2, would reach the intestine intact following ingestion. Furthermore, the unstructured peptic fragments of such lysine rich proteins would contain numerous tryptic cleavage sites, allowing rapid further digestion by trypsin in the intestine, as demonstrated here for the most stable BHL8 peptic fragment. In the unlikely event that some of the BHL proteins did reach the intestine intact, they would probably still be nutritionally available, as both BHL8 and BHL9 were completely or substantially digested within 60 min in SIF and were also devoid of inhibitory activity against intestinal proteases. Unlike the BHL proteins, wild-type CI-2 was completely stable in SIF. Therefore, if any intact wild-type protein reached the intestine, it could pass through the animal unused and inhibit digestive proteases. Previous feeding studies suggested that even this was an unlikely scenario. Rats and mice fed uncooked HiProly barley, a naturally occurring high lysine line with elevated CI-2 levels, gained weight faster than those fed wild-type barley (*17*). Thus, CI-2 did not appear to be a major antinutritional factor, consistent with its lability in SGF evident here.

Prospects for Nutritional Improvement of Plants. Previous studies suggested that proteins with a high degree of stability against proteolysis in vitro were more likely to accumulate to high levels in plants. For example, high methionine mutants of soybean glycinin that were susceptible to proteolysis in vitro did not accumulate in transgenic tobacco, in contrast to the proteolytically more stable wild-type glycinin (*18*). Legumin modified with a methionine rich C-terminal extension was proteolytically unstable in vitro and also did not accumulate in transgenic plants (*19*). In addi-

tion, a high methionine phaseolin mutant that was malfolded (and therefore probably more susceptible to proteases) did not accumulate in transgenic tobacco, in contrast to wild-type phaseolin (*20*). On the basis of their high degree of stability against proteolysis in vitro, BHL8 and BHL9 appeared to be the most likely of the BHL proteins to be expressed at high levels in plants. Considering amino acid composition only, BHL8 may be best suited for poultry, and BHL9 for swine.

The BHL proteins may also be helpful in improving human nutrition in areas where much of the diet comes from a staple cereal crop. The two most limiting amino acids for humans subsisting on maize are lysine and tryptophan (*21*), and these amino acids are greatly enriched in BHL8 and BHL9, relative to maize endosperm protein (Table 1).

The low expression of BHL9 in *E. coli* was surprising, but may not be relevant for predicting expression levels in plants. The stabilizing BHL9 disufide bond may not have formed in vivo in the relatively reducing environment of the *E. coli* cytoplasm, but instead probably formed spontaneously during purification. Disulfide formation would be much more likely to occur in the more oxidizing environment of the secretory pathway in a plant cell.

Versatility of the CI-2 Scaffold. BHL9 and BHL8 now provide two examples of disulfide-stabilized CI-2 derivatives that contain substitutions for over a third of their amino acids and yet retain a high degree of stability in vitro. Thus, the CI-2 platform appears to provide flexibility for solving varying nutritional problems. In proceeding from BHL8 to BHL9, one or more of the amino acid differences resulted in greater thermodynamic stability and also greater proteolytic stability against SIF, trypsin, and chymotrypsin. Although we do not know which of the numerous substitutions contributed to the increased stability in BHL9, it is thought that amino acids with branched side chains, such as threonine and isoleucine, contribute to the thermodynamic stability of proteins by decreasing the entropy of the unfolded form (*22*). The high content of these residues in BHL9 may therefore have contributed to the higher thermodynamic stability, which has been shown with other proteins to be correlated with proteolytic stability (*23*, *24*), as evident here with the SIF data of Table 2.

In conclusion, this study demonstrated that it is possible to achieve two seemingly contradictory goals in proteins engineered for nutritional enhancement—a high degree of stability against proteases in general (and therefore a potentially high expression level in plants) and yet ready digestibility of the proteins by monogastric animals. These two goals were not mutually exclusive for the BHL proteins. In the extremely low pH of the gastric fluid, the proteins were denatured with consequent exposure of proteolytic cleavage sites and rapid digestion. Alternatively, in the more neutral pH that is typical of, for example, the neutral storage protein vacuoles of a plant cell (*25*), the stable BHL proteins such as BHL8 and BHL9 were compactly folded and relatively resistant to proteases. Such compactly folded proteins would also be more likely to escape the ER-associated degradation observed with malfolded proteins (*26*).

ABBREVIATIONS USED

BHL, barley high lysine; CD, circular dichroism; CI-2, chymotrypsin inhibitor-2; GdmCl, guanidinium chloride; SGF, simulated gastric fluid; SIF, simulated intestinal fluid; UV, ultraviolet; WT, wild type.

ACKNOWLEDGMENT

We thank Dr. David Meyer of Pioneer Hi-Bred International for performing the search of the maize databases to identify homologues of barley CI-2 and Dr. Larry Beach for critical review of the manuscript.

LITERATURE CITED

(1) Hans, I. K.; Lee, J. H. The role of synthetic amino acids in monogastric animal production. *Asian-Aus. J. Animal Sci.* **2000**, *13*, 543–560.

(2) Roesler, K. R.; Rao, A. G. A single disulfide bond restores thermodynamic and proteolytic stability to an extensively mutated protein. *Protein Sci.* **2000**, *9*, 1642–1650.

(3) Roesler, K. R.; Rao, A. G. Conformation and stability of barley chymotrypsin inhibitor-2 (CI-2) mutants containing multiple lysine substitutions. *Protein Eng.* **1999**, *12*, 967–973.

(4) Lasztity, R. *The Chemitry of Cereal Proteins*; CRC Press: Boca Raton, FL, 1996.

(5) Board of Trustees. Simulated Gastric Fluid. *The United States Pharmacoepia 23, The National Foundry 18*; The United States Pharmacoepial Convention, Inc.: Rockville, MD, 1995; p 2053.

(6) Astwood, J. D.; Leach, J. N.; Fuchs, R. L. Stability of food allergens to digestion in vitro. *Nat. Biotech.* **1996**, *14*, 1269–1273.

(7) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254.

(8) Scopes, R. K. Measurement of protein by spectrophotometry at 205 nm. *Anal. Biochem.* **1974**, *59*, 277–282.

(9) Pearson, W. R.; Lipman, D. J. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2444–2448.

(10) de Prat Gay, G.; Fersht, A. R. Generation of a family of protein fragments for structure-folding studies. 1. Folding complementation of two fragments of chymotrypsin inhibitor-2 formed by cleavage of its unique methionine residue. *Biochemistry* **1994**, *33*, 7957–7963.

(11) Longstaff, C.; Campbell, A. F.; Fersht, A. R. Recombinant Chymotrypsin Inhibitor 2: Expresion, Kinetic Analysis of Inhibition with a-chymotrypsin and Wild-Type and Mutant Subtilisin BPN′, and Protein Engineering To Investigate Inhibitory Specificity and Mechanism. *Biochemistry* **1990**, *29*, 7339–7347.

(12) Cordero, M. J.; Raventos, D.; Segundo, B. S. Expression of a maize proteinase inhibitor gene in response to wounding and fungal infection: systemic wound-response of a monocot gene. *Plant. J.* **1994**, *6*, 141–150.

(13) Tamayo, M. C.; Rufat, M.; Bravo, J. M.; Segundo, B. S. Accumulation of a maize proteinase inhibitor in response to wounding and insect feeding, and characterization of its activity toward digestive proteinases of *Spodoptera littoralis* larvae. *Planta* **2000**, *211*, 62–71.

(14) Kjaer, M.; Ludvigsen, S.; Sorensen, O. W.; Denys, L. A.; Kindtler, J. et al. Sequence specific assignment of the proton NMR spectrum of barley serine proteinase inhibitor 2. *Carlsberg Res. Commun.* **1987**, *52*, 327–354.

(15) Boisen, S.; Andersen, C. Y.; Hejgaard, J. Inhibitors of chymotrypsin and mircobial serine proteases in barley grains. *Physiol. Plant* **1981**, *52*, 167–176.

(16) Horowitz, M.; Harding, P. E.; Maddox, A. F.; Wishart, J. M.; Akkermans, L. M. A. et al. Gastric and oesophageal emptying in patients with Type 2 (noninsulin-dependent) diabetes mellitus. *Diabetologia* **1989**, *32*, 151–159.

(17) Munck, L.; Karlsson, K. E.; Hagberg, A. Genes for improved nutritional value in barley seed protein. *Science* **1970**, *168*, 985–987.

(18) Jung, R.; Scott, M. P.; Nam, Y. W.; Beaman, T. W.; Bassuner, R. et al. The role of proteolysis in the processing and assembly of 11S seed globulins. *Plant cell* **1998**, *10*, 343−357.
(19) Saalbach, G.; Christov, V.; Jung, R.; Saalbach, I.; Manteuffel, R. et al. Stable expression of vicilin from Vicia faba with eight additional single methionine residues but failure of accumulation of legumin with an attached peptide segment in tobacco seeds. *Mol. Breed.* **1995**, *1*, 245−258.
(20) Pueyo, J. J.; Chrispeels, M. J.; Herman, E. M. Degradation of transport-competent destabilized phaseolin with a signal for retention in the endoplasmic reticulum occurs in the vacuole. *Planta* **1995**, *196*, 586−596.
(21) Glover, D. V.; Mertz, E. T. Nutritional quality of cereal grains: genetic and agronomic improvement; American Society of Agronomy, Inc.: Madison, WI, 1987.
(22) Matthews, B. W. Genetic and Structural Analysis of the Protein Stability Problem. *Biochemistry* **1987**, *26*, 6885−6887.
(23) Liao, H. H. Thermostable mutants of kanamycin nucleotidetransferase are also more stable to proteinase K, urea, detergents, and water-miscible organic solvents. *Enzyme Microb. Technol.* **1993**, *15*, 286−292.
(24) Parsell, D. A.; Sauer, R. T. The structural stability of a protein is an important determinant of its proteolytic susceptibility in *E. coli*. *J. Biol. Chem.* **1989**, *264*, 7590−7595.
(25) Rogers, S. W.; Burks, M.; Rogers, J. C. Monoclonal antibodies to barley aleurain and homologues from other plants. *Plant J.* **1997**, *11*, 1359−1368.
(26) Ellgaard, L.; Molinari, M.; Helenius, A. Setting the standards: Quality control in the secretory pathway. *Science* **1999**, *286*, 1882−1888.
(27) Williamson, M. S.; Forde, J.; Buxton, B.; Kries, M. Nucleotide sequence of barley chymotrypsin inhibitor-2 (CI-2) and its expression in normal and high lysine barley. *Eur. J. Biochem.* **1987**, *165*, 99−106.
(28) McCaldon, P.; Argos, P. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* **1988**, *4*, 99−122.

Received for review February 16, 2001. Revised manuscript received May 8, 2001. Accepted May 9, 2001.

JF010209B